# Multiple-Antenna Systems: From Generic to Hardware-Informed Precoding Designs

A thesis submitted for the degree of Doctor of Philosophy

(Ph.D.)

by

***Ang Li***


Communications and Information Systems Research Group

Department of Electronic and Electrical Engineering

University College London.

April 8, 2018

*To my wife, Xuan*

I, Ang Li, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Sign:

Date:

# Abstract

5G-and-beyond communication systems are expected to be in a heterogeneous form of multiple-antenna cellular base stations (BSs) overlaid with small cells. The fully-digital BS structures can incur significant power consumption and hardware complexity. Moreover, the wireless BSs for small cells usually have strict size constraints, which incur additional hardware effects such as mutual coupling (MC). Consequently, the transmission techniques designed for future wireless communication systems should respect the hardware structures at the BSs. For this reason, in this thesis we extend generic downlink precoding to more advanced hardware-informed transmission techniques for a variety of BS structures.

This thesis firstly extends the vector perturbation (VP) precoding to multiple-modulation scenarios, where existing VP-based techniques are sub-optimal. Subsequently, this thesis focuses on the downlink transmission designs for hardware effects in the form of MC, limited number of radio frequency (RF) chains, and low-precision digital-to-analog converters (DACs). For these scenarios, existing precoding techniques are either sub-optimal or not directly applicable due to the specific hardware constraints. In this context, this thesis first proposes analog-digital (AD) precoding methods for MC exploitation in compact single-user multiple-antenna systems with the concept of constructive interference, and further extends the idea of MC exploitation to multi-user scenarios with a joint optimisation on the precoding matrix and the mutual coupling effect. We further consider precoding for wireless BSs with a limited number of RF chains, in the form of compact

parasitic antenna array as well as hybrid analog-digital structures designed for large-scale multiple-antenna systems. In addition, with a reformulation of the constructive interference, this thesis also considers the low-complexity precoding design for the use of low-resolution DACs for a massive-antenna array at the BSs.

Analytical and numerical results reveal an improved performance of the proposed techniques compared to the state-of-the-art approaches, which validates the effectiveness of the introduced methods.

# Acknowledgements

From the beginning of my Ph.D. study to the construction of this thesis, there are many people that I would like to acknowledge and express my thanks to.

First of all, I would like to give my greatest gratitude to my Ph.D. supervisor, Dr. Christos Masouros, for offering me this great opportunity to pursue my Ph.D. in UCL. From the very beginning, I am deeply encouraged by his attitude and enthusiasm towards research, and I appreciate his generous supports and revisions in every paper we have produced together. It is his valuable guidances and detailed advices that have helped me achieve all my accomplishments, which further lead to this thesis. Honestly, the past three years may be the most important time in my life, and I am happy and grateful that I have spent these years in UCL under his supervision.

Then, I would like to thank my parents for their priceless and unselfish love to me. I deeply appreciate for their encouragement on me to pursue a Ph.D. degree overseas, and I am also grateful for their understanding and support in every decision I have made.

Moreover, I would like to express my special thanks and love to my wife Ms. Xuan Yue. I thank her for your encouragement during my Ph.D. study, which is unique and irreplaceable. I value and cherish every minute we have spent together in London and Leicester, and I treasure the happiness she has brought to my life. I am sure that our happiness will also have a positive effect for our new family member Xinyun.

I would like to further express my appreciation to Prof. Constantinos B.

Papadias, Prof. Mathini Sellathurai, and Prof. A. Lee Swindlehurst for their insightful remarks in the papers we have co-authored. I also appreciate Dr. Ioannis Andreopoulos and Dr. Miguel Rodrigues for attending my MPhil to Ph.D. transfer, and providing helpful suggestions on my research topics and presentation skills. I would also like to appreciate Prof. Kai-Kit Wong and Dr. Hien Quoc Ngo for attending my Ph.D. viva, both of whom have provided insightful comments on my research.

Moreover, I am also thankful to my colleagues in UCL, including but not limited to: Dr. Adrian Garcia Rodriguez, Dr. Pierluigi Vito Amadori, Dr. Lifeng Wang, Mr. Fan Liu, Mr. Pingfan Song, Dr. Meng-Yang Chen, Ms. Jun Qian, Ms. Xiaoyan Hu, Ms. Jialing Liao, Dr. Tongyang Xu, Dr. Ka Lung Law, and Mr. Yin Bi.

Finally, I acknowledge the financial support provided by the China Scholarship Council between 2015-2018.

# Contents

# List of Abbreviations

| | |
|---|---|
| **5G** | Fifth Generation |
| **AD** | Analog-Digital |
| **ADCs** | Analog-to-Digital Converters |
| **AoA** | Angle of Arrival |
| **AoD** | Angle of Departure |
| **AP** | Access Point |
| **AWGN** | Additive White Gaussian Noise |
| **BD** | Block Diagonalisation |
| **BDVP** | Block Diagonalised Vector Perturbation |
| **BER** | Bit Error Rate |
| **BLER** | Block Error Rate |
| **BPSK** | Binary Phase-Shift Keying |
| **BS** | Base station |
| **CI** | Constructive Interference |
| **CSI** | Channel State Information |
| **CVP** | Constructive Vector Perturbation |
| **D2D** | Device-to-Device |
| **DACs** | Digital-to-Analog Converters |

| | |
|---|---|
| **DC** | Direct Current |
| **DoD** | Direction of Departure |
| **DPC** | Dirty-Paper Coding |
| **EbN0** | Energy per bit to Noise ratio |
| **EH** | Energy Harvesting |
| **EM** | ElectroMagnetic |
| **EMF** | ElectroMagnetic-Field |
| **ESPAR** | Electronically Steerable Parasitic Array Radiator |
| **FD** | Fully-Digital |
| **FDD** | Frequency-Division Duplex |
| **ID** | Information Decoder |
| **IoTs** | Internet of Things |
| **JCVP** | Joint Constructive Vector Perturbation |
| **JVP** | Joint Vector Perturbation |
| **KKT** | Karush-Kuhn-Tucker |
| **LoS** | Line-of-Sight |
| **MC** | Mutual Coupling |
| **MF** | Matched Filtering |
| **MIMO** | Multiple-Input Multiple-Output |
| **MISO** | Multiple-Input Single-Output |
| **ML** | Maximum Likelihood |
| **MMSE** | Minimum Mean Squared Error |
| **mmWave** | Millimeter-Wave |
| **MU-MIMO** | Multiuser Multiple-Input Multiple-Output |

| | |
|---|---|
| **MU-MISO** | Multiuser Multiple-Input Single-Output |
| **NLoS** | Non-Line-of-Sight |
| **OFDM** | Orthogonal Frequency-Division Multiplexing |
| **OMP** | Orthogonal Match Pursuit |
| **P2P** | Point-to-Point |
| **PA** | Power Amplifier |
| **PSK** | Phase-Shift Keying |
| **QAM** | Quadrature Amplitude Modulation |
| **QP** | Quadratic Programming |
| **QPSK** | Quadrature Phase-Shift Keying |
| **RF** | Radio Frequency |
| **RZF** | Regularized Zero-Forcing |
| **SDP** | Semi-Definite Programming |
| **SDR** | Semi-Definite Relaxation |
| **SINR** | Signal-to-Interference-plus-Noise Ratio |
| **SISO** | Single-Input Single-Output |
| **SNR** | Signal-to-Noise Ratio |
| **SOCP** | Second-Order Cone Programming |
| **SVD** | Singular Value Decomposition |
| **SWIPT** | Simultaneous Wireless Information and Power Transfer |
| **TDD** | Time-Division Duplex |
| **THP** | Tomlinson-Harashima Precoding |
| **THz** | TeraHertz |
| **UE** | User Equipment |

**ULA** Uniform Linear Array

**UPA** Uniform Planar Array

**UGVP** User Grouping Vector Perturbation

**V-BLAST** Vertical-Bell Laboratories Layered Space-Time

**VP** Vector Perturbation

**ZF** Zero-Forcing

# List of Notations

| | |
|---|---|
| $a$ | Scalar |
| $\mathbf{a}$ | Vector |
| $\mathbf{A}$ | Matrix |
| $j$ | Imaginary unit |
| $\mathscr{O}(\cdot)$ | Order of numerical operations |
| $\mathbb{C}^{m\times n}$ | A $m \times n$ matrix in the complex set |
| $\mathbb{CZ}^{m\times n}$ | A $m \times n$ matrix in the complex integer set |
| $\mathbf{A} \succeq 0$ | Matrix A is postive semi-definite |
| $\mathbb{CN}(\alpha, \beta)$ | Complex normal distribution |
| $\mathbb{N}(\alpha, \beta)$ | Normal distribution |
| $\mathbb{E}\{\cdot\}$ | Expectation of a random variable |
| $\Re(\cdot)$ | Real part of a complex scalar, vector or matrix |
| $\Im(\cdot)$ | Imaginary part of a complex scalar, vector or matrix |
| $(\cdot)^T$ | Transpose |
| $(\cdot)^H$ | Conjugate transpose |
| $\mathrm{tr}\{\cdot\}$ | Trace of a matrix |
| $\det\{\cdot\}$ | The determinant of a square matrix |
| $diag(\mathbf{a})$ | Transformation of the vector $\mathbf{a}$ into a diagonal matrix |
| $\mathrm{vec}(\cdot)$ | Vectorisation operation |

| | |
|---|---|
| $(\cdot)^{-1}$ | Inverse of a square matrix |
| $(\cdot)^{\dagger}$ | Moore-Penrose inverse |
| $\lvert\cdot\rvert$ | Absolute value or modulus |
| $\lVert\cdot\rVert_1$ | 1-norm |
| $\lVert\cdot\rVert_F$ | Frobenius norm |
| $conj(\cdot)$ | Conjugate or element-wise conjugate |
| $[\mathbf{A}]_{m,n}$ | The element in the $m$-th row and $n$-th column of $\mathbf{A}$ |
| $[\mathbf{a}]_m$ | The $k$-th entry in the vector $\mathbf{a}$ |
| $\mathrm{mod}_{\tau}\{\cdot\}$ | Complex modulo operation |
| $\lfloor\cdot\rfloor$ | Floor function |
| $\lceil\cdot\rceil$ | Ceiling function |
| $\arg\{a\}$ | Phase of a complex number $a$ |
| $\mathrm{sort}(\mathbf{a},\mathrm{d})$ | Sort the elements in $\mathbf{a}$ in a descending order |
| $x=\mathrm{find}\{a=\mathbf{b}\}$ | Find function, $[\mathbf{b}]_x=a$ |
| $(a)^{+}$ | $\max\{0,a\}$ |
| $\Gamma(\cdot)$ | Gamma function |
| ${}_2F_1([a,b];c;z)$ | Hypergeometric function |

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Due to the rapid development of wireless applications such as cloud computing, Internet of Things (IoTs) and physical layer security, and their important roles in modern wireless networks, the past few decades have witnessed an exponential growth in both the wireless devices and the wireless data traffic [1]. Furthermore, recent studies in [2] have revealed that a more than 60% annual growth rate can be expected in the next couple of years. Consequently, the increasing need for higher data rates has stimulated both the academia and industry to bring about new techniques, among which multiple-input multiple-output (MIMO) [3, 4] has been widely acknowledged as the most promising one for the past decades. Compared to the single-antenna systems, multiple-antenna systems are able to greatly improve the spectral efficiency via spatial multiplexing [5], and therefore downlink precoding designs that achieve this goal have received extensive research attention.

Currently, most of the existing research on the MIMO systems has been conducted assuming perfect hardware components, where the imperfections in the hardware components have usually been neglected. For example, the mutual coupling (MC) effect [6], which does not exist in single-antenna systems, arises when the inter spacing between the adjacent antenna elements is small compared to the carrier wavelength. The negligence of the MC effect will usually lead to a performance degradation for MIMO systems

in practical deployments [7], especially for size-constrained small-scale base stations (BSs) towards the fifth generation (5G) and future wireless communications. Nevertheless, the hardware imperfection in the form of MC can also be beneficial. For example, the compact parasitic-antenna systems can exploit the MC effect to form the desired radiation pattern by the use of tunable antenna loads, where the total number of required radio frequency (RF) chains at the BSs is reduced [8].

On the other hand, the large-scale or equivalently massive MIMO introduced in [9, 10] has also rapidly risen as one of the most promising techniques for the future 5G communications. Conceptually, massive MIMO is the same as the generic multiple-antenna systems, with the number of antennas scaled up to the order of hundreds or even thousands [9]. Compared to small-scale MIMO, massive MIMO systems are able to offer unprecedented benefits, such as extremely high throughputs and extremely low bit error rates (BERs) [10]. While the benefits from massive MIMO systems are appealing, the practical implementation of massive MIMO is challenging due to the prohibitive hardware complexity [9]. Indeed, for hardware-ideal MIMO BSs, each transmit antenna element needs to be connected to one dedicated RF chain with a pair of infinite-resolution[1] digital-to-analog converters (DACs). Employing this approach for large-scale MIMO systems will enforce the massive MIMO BSs to be equipped with an equally large number of RF chains and pairs of high-resolution DACs. This requirement, however, significantly increases the hardware costs and more importantly the consequent power consumption for massive MIMO BSs, which is both cost-inefficient and power-inefficient. For this reason, it is more preferable to simplify the hardware structures in practical massive MIMO systems, where the corresponding hardware costs are reduced, and a balance between performance and hardware complexity is achieved [11].

Because of the considerations mentioned above, for small-scale compact

[1] In practical small-scale MIMO systems, high-resolution (usually more than 8 bits) DACs are employed.

MIMO BSs, recent research has focused on the techniques that reduce the MC effect [12, 13]. In addition, the parasitic-antenna systems that can further benefit from the MC effect have also been extensively studied [8, 14, 15]. For the areas of massive MIMO, the hybrid analog-digital (AD) structures with a reduced number of RF chains [11] and the employment of low-cost finite-precision DACs [16] have both been introduced to balance performance, hardware complexity and the consequent power consumption, where both techniques have received increasing research attention. Given the great necessity of advanced techniques for a variety of wireless BS structures, this thesis focuses on the designs of innovative and hardware-informed transmission strategies for both small-scale and large-scale multiple-antenna BSs.

## 1.1 Aim and Motivation

While precoding approaches have been widely studied for multiple-antenna systems, there still exist numbers of unresolved areas that require further investigations, especially for size-constrained compact multiple-antenna BSs and large-scale MIMO systems.

While it is well established that non-linear vector perturbation (VP) precoding achieves a significant performance gain over linear precoding schemes [17], generic VP approaches are not applicable to multiple-modulation scenarios where multiple modulation types are employed, because the modulo basis for the sphere search is modulation-dependent. In the existing literature, however, the introduced block-diagonalised VP (BDVP) and user-grouping VP (UGVP) precoding approaches for multiple-modulation cases [18] can only achieve a sub-optimal performance.

With respect to the MC effect that exists in compact multiple-antenna BSs, the compensation methods for the MC effect have been extensively studied in [7, 19, 20] for generic fully-digital (FD) compact MIMO systems. Nevertheless, these proposed approaches are from the perspective of antenna design instead of a signal processing perspective, and thus are only limited

to specific antenna arrays or carrier frequencies. On the other hand, instead of compensating for this effect, there already exist studies on the exploitation of the MC effect. Innovative compact parasitic-antenna systems, also known as electronically steerable parasitic array radiators (ESPARs), were firstly proposed in [14] to exploit the MC effect instead of avoiding this effect. The transmission schemes for ESPARs in the beam-space domain have been designed in [8, 15], while more recent studies extend the downlink precoding designs for conventional MIMO systems to ESPARs by mapping the precoded signals to the currents for the parasitic-antenna array [21, 22]. However, in practical implementation ESPARs can only employ tunable loads with a finite precision. This is critical for the applications of ESPARs, as the transmit signals are dependent on the values of the tunable antenna loads, and the quantisation in the tunable loads may lead to a mismatch between the desired and actual radiation pattern for ESPARs, which requires further investigations. Moreover, it is still not clear in the existing literature whether the exploitation of the MC effect by the use of tunable antenna loads in the case of ESPARs can be extended to the conventional FD MIMO BSs to achieve further performance gains.

For massive MIMO BSs, while the hybrid AD structures have been extensively studied [11, 23, 24], most of the contributions have focused on the reduction in the hardware complexity, while ignoring their advantages in the energy efficiency. With a reduced number of RF chains, the power consumption at the BSs is greatly reduced, which achieves an energy-efficient transmission. Particularly, this aspect is not limited to large-scale multiple-antenna systems and can be extended to small-scale MIMO systems, which motivates our study in energy-efficient MIMO SWIPT systems with hybrid precoding. On the other hand, in the case of employing limited-precision analog devices, most of the studies have focused on the performance analyses with low-resolution analog-to-digital converters (ADCs) in the uplink [25, 26]. For precoding techniques in the massive MIMO downlink with

finite-precision DACs, there are only limited contributions in the existing literature [16, 27, 28]. Moreover, these existing methods either suffer from a significant performance loss compared to the ideal unquantised case [16, 27], or include computationally-inefficient iterative algorithms [28]. This then motivates the development of novel and more advanced precoding techniques that can achieve a more favourable performance-complexity tradeoff.

## 1.2 Main Contributions

This thesis aims to compensate for the performance losses introduced by the use of imperfect hardware components, and further exploit the specific hardware structures at the BSs to achieve an improved performance-complexity tradeoff through the introduction of novel hardware-informed downlink transmission methods. The main contributions of this thesis can be summarised as follows:

- Introduction of both a joint VP (JVP) precoding and a joint constructive VP (JCVP) for multiple-modulation scenarios (Chapter 3). The proposed JVP approach is shown to achieve a comparable performance to conventional VP precoding in multiple-modulation cases, where existing approaches achieve an inferior performance, while the proposed JCVP method is shown to achieve a similar performance with a reduced computational cost.

- Exploitation of the MC effect for conventional FD compact point-to-point (P2P) MIMO systems with tunable antenna loads via an optimisation on the loads based on the constructive interference (CI) formulation (Chapter 4). Numerical results reveal that further gains can be obtained through the exploitation of the adaptive MC effect in compact P2P MIMO systems with tunable loads.

- Extension of the MC exploitation to multiuser multiple-input single-output (MU-MISO) systems in the downlink by the proposition of a

joint AD precoding for signal-to-interference-plus-noise ratio (SINR) balancing optimisation (Chapter 5). It is shown by simulation validation that the use of tunable antenna loads can achieve an improved performance for both perfect channel state information (CSI) and imperfect CSI.

- Performance analyses for ESPARs in the presence of quantised antenna loads with imperfect CSI, and design of a quantisation-robust downlink transmission approach based on convex optimisation (Chapter 6). The analytical results are shown to match the simulated results, and the proposed robust method is shown to be capable of greatly alleviating the performance loss by the quantisation in the antenna loads.

- Introduction of a hybrid precoding design via a virtual path selection approach for millimeter-wave (mmWave) communications, and the application of hybrid precoding for the energy-efficient simultaneous wireless information and power transfer (SWIPT) systems via a geometrical approach (Chapter 7). A close-to-optimal performance is observed for the method based on virtual path selection, while the improved energy efficiency is observed by the proposed algorithm for the considered MIMO SWIPT systems.

- Design of a low-complexity symbol-scaling method for massive MIMO downlink with 1-bit DACs, based on the CI formulation and coordinate transformation (Chapter 8). The improved performance-complexity tradeoff of the proposed symbol scaling approach compared to existing schemes in the literature is validated by extensive numerical results.

## 1.3 Thesis Organisation

Subsequent to this introductory chapter, this thesis is organised following the structure depicted in Fig. 1.1 and described in the sequel.

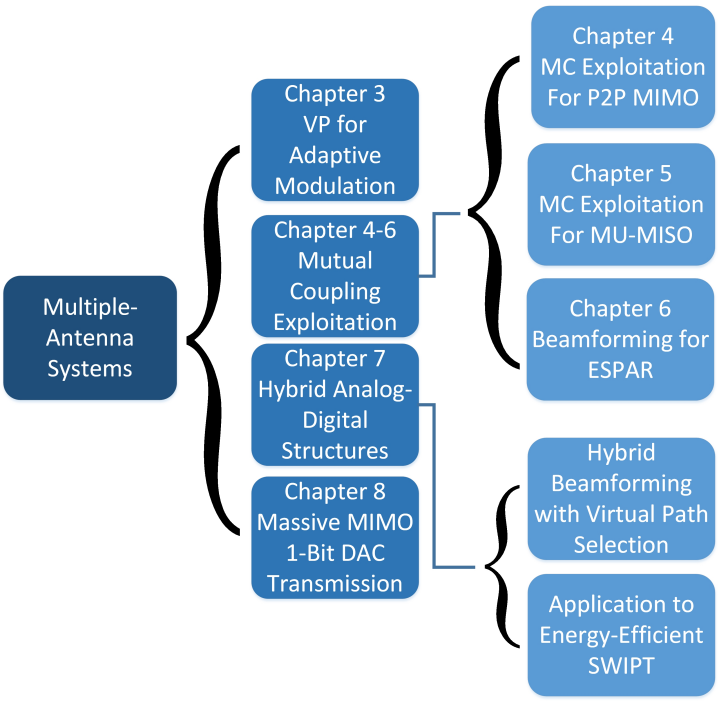



**Figure 1.1:** Thesis Organisation

Chapter 2 provides a review of multiple-antenna systems, which constitutes the basis of this thesis. In particular, this chapter focuses on the state-of-the-art precoding techniques in the downlink. The introduction of two typical hybrid BS structures that require a reduced number of RF chains is also included.

Chapter 3 first introduces the proposed JVP precoding scheme for multiple-modulation scenarios. The proposed method enables a joint sphere search for the perturbation vectors for multiple modulations via a constellation scaling approach. Building upon the proposed JVP approach and selective perturbation methods, a JCVP precoding is further introduced for both phase-shifting keying (PSK) and multiple-modulation cases based on the CI formulation, where a joint sphere-search and symbol-scaling approach is proposed.

Chapter 4 presents a CI-based precoding that exploits the MC effect for compact P2P MIMO systems with tunable antenna loads. It is firstly analytically proven that the MC effect cannot be fully eliminated by solely tun-

ing the load values, followed by the introduction of the proposed precoding method based on convex optimisation. The proposed approach manipulates the load values on each antenna element, such that the interference introduced by the MC effect aligns constructively to the useful signal vector. The practical implementation of the proposed technique is further included and discussed in detail.

Chapter 5 extends the exploitation of the MC effect to the multi-user case and proposes a joint AD precoding for the downlink transmission. A joint optimisation is firstly formulated with a SINR balancing optimisation in the digital domain and an optimisation of the tunable antenna loads in the analog domain, followed by the introduction of a low-complexity decoupled approach. Subsequently, the case for quantised analog loads are considered and a sequential search scheme is proposed based on a greedy algorithm to efficiently obtain the desired load values. The robust designs for imperfect CSI are also included.

Chapter 6 considers the downlink transmission for ESPARs in the presence of quantised loads and imperfect CSI. Mathematical analyses on the performance of ESPARs with quantised loads are firstly performed for both perfect CSI and imperfect CSI. A quantisation-robust downlink transmission strategy that minimises the Euclidean distance between the desired and actual current vector for the ESPAR array is further proposed to alleviate the mismatch by quantisation in the tunable antenna loads.

Chapter 7 presents a computationally efficient hybrid precoding design via a virtual path selection method, where three distinct design criteria are introduced and each achieves a different performance-complexity tradeoff. The hybrid precoding is further extended to the case of MIMO SWIPT for the purpose of an energy-efficient transmission, where both a low-complexity approach and an iterative algorithm based on a geometrical representation are proposed.

Chapter 8 proposes downlink transmission methods for massive MIMO

downlink with hardware imperfections in the form of 1-bit DACs. Based on the CI formulation, a quantised linear approach is first considered, where the optimal precoding matrix is obtained. A non-linear mapping method is further considered where we directly design the transmit signal vector. A two-step relaxation-normalisation process is adopted to solve the non-convex optimisation. A low-complexity three-stage symbol-scaling scheme is further proposed, where the quantised transmit signal on each antenna element is selected sequentially.

Chapter 9 concludes this thesis with a summary of the contributions in the previous chapters, and discusses possible future extensions of the research within the frame of this thesis.

## 1.4 List of Publications

The above contributions in this thesis have led to the following publications.

**Journal Papers:**

[J1] **A. Li**, C. Masouros, F. Liu, and A. L. Swindlehurst, "Massive MIMO 1-Bit DAC Transmission: A Low-Complexity Symbol Scaling Approach", *IEEE Transactions on Wireless Communications*, available online: http://arxiv.org/abs/1709.08278, review after major revision.

[J2] **A. Li**, and C. Masouros, "Energy-Efficient SWIPT: From Fully-Digital to Hybrid Analog-Digital Beamforming", *IEEE Transactions on Vehicular Technology*, DOI:10.1109/TVT.2017.2782775, accepted, 2017.

[J3] **A. Li**, C. Masouros, and M. Sellathurai, "Analog-Digital Beamforming in the MU-MISO Downlink by use of Tunable Antenna Loads", *IEEE Transactions on Vehicular Technology*, DOI:10.1109/TVT.2017.2776563, accepted, 2017.

[J4] **A. Li**, C. Masouros, and C. B. Papadias, "MIMO Transmission for Single-fed ESPAR with Quantized Loads", *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 2863-2876, July 2017.

[J5] **A. Li**, and C. Masouros, "Exploiting Constructive Mutual Coupling in P2P MIMO by Analog-Digital Phase Alignment", *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1948-1962, Mar. 2017.

[J6] **A. Li**, and C. Masouros, "Hybrid Analog-Digital Millimeter-Wave MU-MIMO Transmission with Virtual Path Selection", *IEEE Communications Letters*, vol. 21, no. 2, pp. 438-441, Feb. 2017.

[J7] **A. Li**, and C. Masouros, "A Two-stage Vector Perturbation Scheme for Adaptive Modulation in Downlink MU-MIMO", *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 7785-7791, Sept. 2016.

[J8] **A. Li**, and C. Masouros, "A Constellation Scaling Approach to Vector Perturbation in Downlink MU-MIMO", *IEEE Wireless Communications Letters*, vol. 4, no. 3, pp. 289-292, Mar. 2015.

### Conference Papers:

[C1] **A. Li**, C. Masouros, and A. L. Swindlehurst, "1-Bit Massive MIMO Downlink Based on Constructive Interference", submitted to *26th European Signal Processing Conference (EUSIPCO) 2018*.

[C2] **A. Li**, C. Masouros, and F. Liu, "Hybrid Precoding with 1-Bit DAC Transmission for Interference Exploitation", submitted to *26th European Signal Processing Conference (EUSIPCO) 2018*, ***invited paper***.

[C3] **A. Li**, and C. Masouros, "Hybrid Massive MIMO Unlicensed Transmission with 1-Bit Quantization", *2017 IEEE Global Communications Conference (GLOBECOM) Workshops*, *accepted*, Singapore, 2017.

[C4] **A. Li**, C. Masouros, M. Sellathurai, and C. B. Papadias, "Tunable Load MIMO with Quantized Loads", *25th European Signal Processing Conference (EUSIPCO) 2017*, pp. 1-6, Kos, Greece, 2017.

[C5] **A. Li**, C. Masouros, and M. Sellathurai, "Analog-Digital Beamforming for MIMO with Tunable Loads by Mutual Coupling Exploitation", *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1-5, Sapporo, Japan, 2017.

[C6] **A. Li**, and C. Masouros, "Energy Efficient MIMO SWIPT by Hybrid Analog-Digital Beamforming", *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1-5, Sapporo, Japan, 2017.

[C7] **A. Li**, and C. Masouros, "Mutual Coupling Exploitation for Point-to-point MIMO by Constructive Interference", *2017 IEEE International Conference on Communications (ICC)*, pp. 1-6, Paris, France, 2017.

[C8] **A. Li**, and C. Masouros, "Hybrid Precoding and Combining Design for Millimeter-Wave Multi-User MIMO based on SVD", *2017 IEEE International Conference on Communications (ICC)*, pp. 1-6, Paris, France, 2017.

[C9] **A. Li**, and C. Masouros, "Exploiting Mutual Coupling by Means of Analog-Digital Zero Forcing", *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3564-3568, New Orleans, LA, 2017.

[C10] **A. Li**, and C. Masouros, "Performance Analysis for Single-fed ESPAR in the Presence of Impedance Errors and Imperfect CSI", *2016 IEEE International Conference on Communications (ICC)*, pp. 1-6, Kuala Lumpur, 2016.

### Contributions during Ph.D. that are not included in this thesis:

### Journal Papers:

1. **A. Li**, and C. Masouros, "Interference Exploitation Precoding Made Practical: Closed-Form Solutions with Optimal Performance", *IEEE Transactions on Wireless Communications*, available online: https://arxiv.org/abs/1712.07846, major revision.
2. F. Liu, L. Zhou, C. Masouros, **A. Li**, W. Luo, and A. Petropulu, "Towards Dual-Functional Radar-Communication Systems: Optimal Waveform Design", *IEEE Transactions on Signal Processing*, available online: https://arxiv.org/abs/1711.05220, under review.
3. F. Liu, C. Masouros, **A. Li**, H. Sun, and L. Hanzo, "MU-MIMO Communications with MIMO Radar: From Co-existence to Joint Transmission", *IEEE Transactions on Wireless Communications*, DOI:10.1109/TWC.2018.2803045, accepted, 2018.
4. F. Liu, C. Masouros, **A. Li**, T. Ratnarajah, and J. Zhou,"Interference Exploitation for Radar and Cellular Coexistence: The Power-Efficient Approach", *IEEE Transactions on Signal Processing*, available online: https://arxiv.org/abs/1704.08920, under review.
5. F. Liu, C. Masouros, **A. Li**, and T. Ratnarajah, "Robust MIMO Beamforming for Cellular and Radar Coexistence", *IEEE Wireless Communications Letters*, vol. 6, no. 3, pp. 374-377, Apr. 2017.

### Conference Papers:

1. **A. Li**, and C. Masouros, "Closed-Form Constructive Interference Precoding for the MU-MIMO Downlink", submitted to *2018 IEEE Global Communications Conference (GLOBECOM)*.
2. L. Zhou, F. Liu, C. Masouros, **A. Li**, W. Jiang, and W. Luo, "Optimal Waveform Design for Dual-Functional MIMO Radar-Communication Systems", submitted to *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*.
3. F. Liu, L. Zhou, C. Masouros, **A. Li**, J. Zhou, and W. Luo, "Optimal MIMO Radar-Communication Waveform Design under Constant Modulus Constraint", submitted to *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, ***invited paper***.
4. **A. Li**, and C. Masouros, "Constructive Interference Beamforming for Cooperative Dual-Hop MIMO Relay Systems", *2018 IEEE 87th Vehicular Technology Conference (VTC-Spring)*, ***invited paper***, *accepted*.

5. F. Liu, C. Masouros, **A. Li**, J. Zhou, and L. Hanzo, "Simultaneous Target Detection and Multi-user Communications Enabled by Joint Beamforming", *IEEE Radar Conference 2018*, accepted.
6. F. Liu, C. Masouros, **A. Li**, and T. Ratnarajah, "Radar and Communication Coexistence Enabled by Interference Exploitation", *2017 IEEE Global Communications Conference (GLOBECOM), accepted*, Singapore, 2017.

# Chapter 2

# Multiple-Antenna Wireless Communication Systems

This chapter introduces fundamental concepts and techniques with respect to multiple-antenna systems that are relevant to this thesis, which include channel models, MC modelling and imperfect CSI models. In particular, closed-form precoding approaches as well as precoding based on optimisation are presented to compare their potential benefits and main drawbacks. In addition, we briefly describe two typical hybrid AD precoding structures at the BS. Throughout this thesis, we discuss narrowband single-carrier multiple-antenna systems.

## 2.1 MIMO Communications - Fundamentals

Due to the increasing demand for higher data rates and reliability for wireless networks, MIMO techniques have appeared and received extensive research attention [3, 5]. To support spatial multiplexing, parallel data streams can be transmitted simultaneously with multiple antennas deployed at the BS. To improve the reliability, space-time coding techniques can be employed by sending copies of the same information across the antenna array [32].

In a wireless P2P MIMO system, as depicted in Fig. 2.1, the data symbol vector is denoted as $\mathbf{s}$, and one BS with $N_t$ antennas transmits wireless signals to a single receiver with a total number of $N_r$ receive antennas. Math-

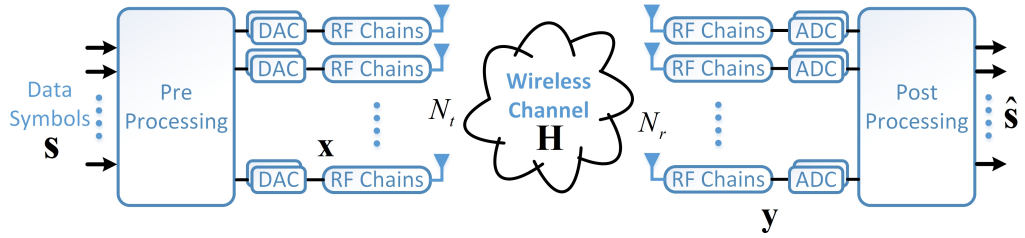



**Figure 2.1:** A block diagram of point-to-point MIMO systems

ematically, the signal vector at the receiver can be expressed as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N_r} \end{bmatrix} = \begin{bmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,N_t} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,N_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r,1} & h_{N_r,2} & \cdots & h_{N_r,N_t} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_t} \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{N_r} \end{bmatrix}, \tag{2.1}$$

where each $h_{i,j}$ represents the complex channel gain from the $j$-th transmit antenna to the $i$-th receive antenna, which will be discussed in detail in Chapter 2.3. Each $x_m$ and $y_n$ denotes the transmit and received signal on the $m$-th transmit antenna and $n$-th receive antenna, respectively. Each $n_k$ represents the additive white Gaussian noise (AWGN) at the $k$-th receive antenna with zero mean and variance $\sigma^2$. Equivalently, (2.1) can be expressed in a matrix form as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \tag{2.2}$$

where $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ and $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ denote the transmit and receive signal vector, respectively. $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix, and $\mathbf{n} \in \mathbb{C}^{N_r \times 1}$ denotes the noise vector, where equivalently we obtain $\mathbf{n} \sim \mathbb{CN}\left(\mathbf{0}, \sigma^2 \cdot \mathbf{I}\right)$. When a transmit pre-processing is employed for the data symbol vector $\mathbf{s}$, we express $\mathbf{x} = \mathbf{T}\mathbf{s}$, where $\mathbf{T}$ is the transmit precoding matrix. At the receiver, a post-processing is employed on the received signal vector $\mathbf{y}$, and the symbol vector ready for demodulation can be expressed as $\mathbf{r} = \mathbf{R}^H\mathbf{y}$.

**Capacity:** When perfect knowledge of the channel $\mathbf{H}$ is known at both the transmitter and receiver, we can decompose the MIMO channel into sev-

eral parallel and non-interfering single-input single-output (SISO) channels, by the use of the singular value decomposition (SVD) [3]. Accordingly, for a given maximum available transmit power, the capacity of the MIMO channel can be obtained by a water-filling algorithm [3], i.e., more power is allocated to better-conditioned wireless channels, while less or even zero power is allocated to worse-conditioned wireless channels.

**Diversity Gain:** Diversity gain defines the improvement in the reliability of the communication links [4, 32]. By receiving multiple (ideally independent) copies of identical transmitted signals, the detection performance can be improved based on combination techniques. In fact, the more independent fading channels and copies, the faster BER drops as a function of the transmit signal-to-noise ratio (SNR). Mathematically, the diversity gain is defined as

$$d = -\lim_{\rho\to\infty} \frac{\log\{P_e(\rho)\}}{\log\{\rho\}}, \tag{2.3}$$

where $\rho$ denotes the transmit SNR, and $P_e(\rho)$ denotes the corresponding error probability for a specific value of $\rho$.

**Multiplexing Gain:** The use of multiple antennas at the transmitters and receivers can create an additional spatial dimension for communications, which is known as a degree of freedom gain or multiplexing gain [5]. This gain can be exploited by simultaneously transmitting multiple parallel data streams, leading to an increase in the data rates with no additional power invested. The multiplexing gain characterises the data rate improvement by the deployment of multiple antennas, which is mathematically defined as

$$m = \lim_{\rho\to\infty} \frac{\log\{R(\rho)\}}{\log\{\rho\}}, \tag{2.4}$$

where $R(\rho)$ denotes the data rates at a given SNR value. There exists a fundamental tradeoff between the diversity gain and the multiplexing gain, which is described in [5] from an information theoretic perspective.

**Bit Error Rate (BER)** is the performance metric employed in most chap-

ters of this thesis. When the hardware imperfections in the following chapters are considered, the Shannon capacity formula may not be valid, and subsequently the BER becomes the most useful metric for physical layer techniques. BER is defined as the ratio of the erroneous bits to the total number of transmitted bits, given by

$$P_b = \frac{N_e}{N_b}, \tag{2.5}$$

where $N_e$ and $N_b$ denotes the number of erroneous bits and total transmitted bits, respectively. To decode the transmit signal vector $\mathbf{x}$, an equaliser $\mathbf{W} \in \mathbb{C}^{N_r \times N_t}$ is required at the receiver to manage the inter-symbol interference. Specifically, when we assume perfect CSI at the receiver side and the zero-forcing (ZF) combining is employed, the equalisation matrix $\mathbf{W}$ is given by

$$\mathbf{W} = \left(\mathbf{H}^H \mathbf{H}\right)^{-1} \mathbf{H}^H, \; N_r \geq N_t. \tag{2.6}$$

For PSK modulations in Rayleigh fading channels that are introduced in Chapter 2.3.1 and Chapter 2.3.2, the corresponding BER expression can be obtained in an analytical form, expressed as

$$P_b = \sum_{k=1}^{N_t} \left\{ \frac{1}{2} - \sqrt{\frac{\gamma_k}{\pi}} \cdot \frac{\Gamma\left(N_r - N_t + \frac{3}{2}\right)}{\Gamma\left(N_r - N_t + 1\right)} \times {}_2F_1\left( \left[ \frac{1}{2}, N_r - N_t + \frac{3}{2} \right]; \frac{3}{2}; -\gamma_k \right) \right\}, \tag{2.7}$$

where $\Gamma(\cdot)$ is the Gamma function, and ${}_2F_1\left([a,b];c;z\right)$ is the hypergeometric function [29]. We denote the covariance matrix for the channel vector of user $k$ as $\mathbf{C}_{\mathbf{h}_k^T}$. When $\mathbf{C}_{\mathbf{h}_k^T}$ for each user $k$ is identical, we further introduce the notation $\mathbf{C}_{\mathbf{H}}$. $\gamma_k$ can then be obtained based on the SNR per bit, a.k.a., $\frac{E_b}{N_0}$ and the covariance matrix $\mathbf{C}_{\mathbf{H}}$, which is obtained as

$$\gamma_k = \frac{1}{\log_2(\mathbb{N}) \cdot \sigma^2 \left[\mathbf{C}_{\mathbf{H}}^{-1}\right]_{k,k}} \tag{2.8}$$

for $\mathbb{N}$-PSK. In (2.8), $\left[\mathbf{C}_{\mathbf{H}}^{-1}\right]_{k,k}$ is the $k$-th diagonal entry in the inverse of the covariance matrix $\mathbf{C}_{\mathbf{H}}$ as the covariance matrix. The detailed derivation for

(2.7) can be found in [30], and [31] discusses the relationship between the ZF and MMSE equalisers from an information theoretic perspective.

## 2.2 Downlink Transmission Methods

Compared to single-antenna systems, multiple-antenna systems require additional signal processing techniques at the transmitter, receiver, or at both sides. Dependent on which side the processing is employed at, the signal processing can be divided into detection techniques at the receiver side and precoding techniques at the transmitter side, respectively. In a multi-user MIMO system, due to the separation in physical spaces, it is generally difficult to perform a joint processing (detection) of the data streams for different users in the downlink. Therefore, receive combining methods are usually employed in the uplink at the BS side, while precoding is preferred in the downlink transmission. In addition, with the knowledge of the channel at the BS, the use of precoding can transfer the signal processing process from the user side to the BS to further alleviate the computational burden of the users, which makes precoding a popular technique.

We move to consider a multi-user MIMO system, as depicted in Fig. 2.2, where the BS with $N_t$ transmit antennas communicates with a total number of $K$ single-antenna users (a MU-MISO system). In the considered scenario, we decompose the channel matrix $\mathbf{H}$ into

$$\mathbf{H} = \left[\mathbf{h}_1^T, \mathbf{h}_2^T, \cdots, \mathbf{h}_K^T\right]^T, \tag{2.9}$$

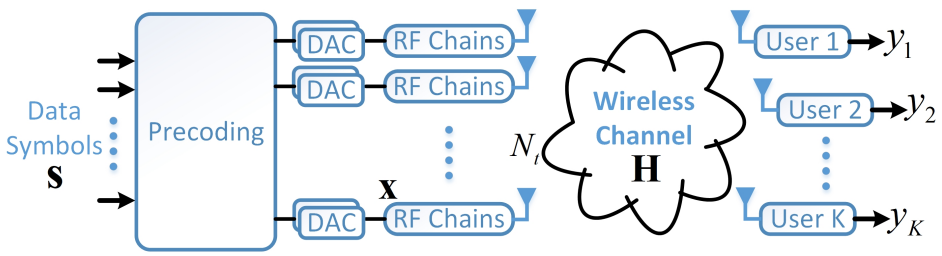



**Figure 2.2:** A block diagram of multi-user MIMO systems

and each $\mathbf{h}_k \in \mathbb{C}^{1\times N_t}$ represents the channel vector from the BS to the $k$-th user. The received signal for the $k$-th user can be expressed as

$$y_k = \mathbf{h}_k \mathbf{x} + n_k, \ \ \forall k \in \{1, 2, \cdots, K\}. \tag{2.10}$$

For most of the precoding techniques, the received signal $y_k$ is ready for demodulation and no additional signal processing is required at the receiver side.

As this thesis focuses on the designs of downlink transmission approaches, existing precoding methods are briefly reviewed in this section. Generally, transmit processing schemes can be divided into closed-form precoding approaches and optimisation-based precoding methods, where closed-form precoding can be further classified into linear and non-linear approaches. In general, linear precoding approaches aim to cancel the multi-user interference based on the knowledge of the channel, which usually require a relatively low computational cost [3]. The non-linear precoding methods can achieve additional performance gains over linear approaches, at the cost of increased computational burdens [17, 18, 34, 35, 36, 37, 38]. In addition, popular optimisation-based downlink precoding methods are also presented, including power minimisation [39], SINR balancing [40], and CI-based optimisation [41].

### 2.2.1 Linear Precoding

Linear precoding represents a set of simple transmission approaches, where the precoded signal vector $\mathbf{x}$ is a linear combination of the data symbols to be transmitted. Mathematically, this is expressed as

$$\mathbf{x} = \mathbf{P}\mathbf{s} = \frac{1}{f} \cdot \mathbf{W}\mathbf{s}. \tag{2.11}$$

In (2.11), $\mathbf{P} = \frac{1}{f} \cdot \mathbf{W} \in \mathbb{C}^{N_t \times K}$ is the precoding matrix, where $f$ represents a scaling factor to ensure that the power of the precoded signals $\mathbf{x}$ is constrained,

given by[1]

$$f = \|\mathbf{W}\|_F = \sqrt{\mathrm{tr}\{\mathbf{W}\mathbf{W}^H\}}, \tag{2.12}$$

which is also known as the amplification factor. In the following we present typical linear precoding methods, which have closed-form expressions and are generally computationally efficient for multiple-antenna systems.

**Matched Filtering (MF)** is the simplest precoding method in the literature, which maximises the received SNR while ignoring the multi-user interference. The precoding matrix is formulated as the Hermitian of the channel matrix $\mathbf{H}$, given by [42]

$$\mathbf{P}_{MF} = \frac{1}{f_{MF}} \cdot \mathbf{H}^H = \frac{\mathbf{H}^H}{\sqrt{\mathrm{tr}\{\mathbf{H}\mathbf{H}^H\}}}. \tag{2.13}$$

MF precoding can achieve a promising performance in noise-limited scenarios (low SNR regimes or large-scale MIMO scenarios), while its performance is significantly degraded in interference-limited scenarios.

**Zero-Forcing (ZF)** precoding is also a simple precoding method that has been extensively studied [43, 44]. This precoding approach forces the multi-user interference to be zero for each user by employing a Moore-Penrose inverse of the wireless channel $\mathbf{H}$, where the precoding matrix is given by

$$\mathbf{P}_{ZF} = \frac{1}{f_{ZF}} \cdot \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} = \frac{\mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1}}{\sqrt{\mathrm{tr}\left\{\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\right\}}},\ N_t \geq K. \tag{2.14}$$

ZF precoding is shown to achieve an improved performance over MF in the high SNR regime.

**Regularised Zero-Forcing (RZF)** requires a similar computational cost compared to ZF precoding, while achieves an improved performance over ZF by introducing a regularisation factor to deal with ill-conditioned channel

[1]This is a short-term power constraint, and a long-term power constraint can also be applied by expressing $f = \mathbb{E}\left\{\sqrt{\mathrm{tr}\{\mathbf{W}\mathbf{W}^H\}}\right\}$. It is shown that the performance difference is very small [33].

matrices $\mathbf{H}$ [33]. The precoding matrix for RZF precoding is given by

$$\mathbf{P}_{RZF} = \frac{1}{f_{RZF}} \cdot \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \alpha \cdot \mathbf{I}\right)^{-1} = \frac{\mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \alpha \cdot \mathbf{I}\right)^{-1}}{\sqrt{\mathrm{tr}\left\{ (\mathbf{H}\mathbf{H}^H + \alpha \cdot \mathbf{I})^{-1} \mathbf{H}\mathbf{H}^H (\mathbf{H}\mathbf{H}^H + \alpha \cdot \mathbf{I})^{-1} \right\}}}, \tag{2.15}$$

where the optimal regularisation factor is $\alpha = K\sigma^2$ for the considered system [33].

### 2.2.2 Non-Linear Precoding

Compared to the linear precoding methods mentioned above, non-linear precoding approaches usually require more sophisticated signal processing techniques at the transmitter side, and may also incur additional processing at the receiver side. For non-linear methods, the precoded signals are no longer a linear combination of the data symbols. The most common non-linear precoders in the existing literature are briefly included in the following.

**Dirty Paper Coding (DPC)** is able to achieve the capacity of the MIMO channel [45]. Nevertheless, the practical implementation of DPC is difficult, because the capacity-achieving DPC assumes an infinite length of codewords and includes complicated sphere search algorithms.

**Tomlinson-Harashima Precoding (THP)** performs a successive subtraction of the known interference at the BS prior to data transmission [46]. THP can be viewed as an application of the Vertical-Bell Laboratories Layered Space-Time (V-BLAST) detection algorithm [47] at the transmitter side. To be more specific, THP first performs a decomposition of the channel matrix

$$\mathbf{H} = \mathbf{L}\mathbf{F}^H \tag{2.16}$$

as the multiplication of a lower-triangle matrix $\mathbf{L}$ and a unitary matrix $\mathbf{F}$. The precoded signal vector for THP is then expressed as

$$\mathbf{x}_{THP} = \mathbf{F}\tilde{\mathbf{x}}_{THP}, \tag{2.17}$$

where $\tilde{\mathbf{x}}_{THP}$ is obtained based on the data symbols, given by

$$[\tilde{\mathbf{x}}_{THP}]_k = \mathrm{mod}_\tau \left\{ s_k - \sum_{l=1}^{k-1} [\mathbf{B}]_{k,l} [\tilde{\mathbf{x}}_{THP}]_l \right\}, \ \forall k \in \{1, 2, \cdots, K\}. \tag{2.18}$$

In (2.18), $\mathrm{mod}_\tau \{x\}$ represents a complex modulo function that is performed for the real and imaginary part of $x$ independently, given by

$$\mathrm{mod}_\tau \{x\} = \left( \Re(x) - \tau \cdot \left\lfloor \frac{\Re(x) + \tau/2}{\tau} \right\rfloor \right) + j \left( \Im(x) - \tau \cdot \left\lfloor \frac{\Im(x) + \tau/2}{\tau} \right\rfloor \right), \tag{2.19}$$

where $\tau$ denotes the modulo basis and $\lfloor \cdot \rfloor$ is the floor function. The matrix $\mathbf{B}$ represents the effective THP channel obtained as

$$\mathbf{B} = \mathbf{GHF}, \tag{2.20}$$

where $\mathbf{G}$ is a diagonal matrix that represents a complex scaling at each receiver. For the THP approach, each diagonal entry in $\mathbf{G}$ is equal to the inverse of the corresponding diagonal entry in $\mathbf{L}$, i.e.,

$$g_k = [\mathbf{G}]_{k,k} = \frac{1}{[\mathbf{L}]_{k,k}}. \tag{2.21}$$

At each receiver, in addition to the scaling operation $g_k$, a modulo operation is also required before demodulation [47].

**Vector Perturbation (VP) Precoding** is the focus of Chapter 3 in this thesis and is a modification of ZF precoding. Upon ZF, VP precoding further performs a perturbation on the data symbols before transmission, such that the resulting transmit signals are better aligned to the eigenvectors of the channel inverse matrix, which leads to a significant decrease in the scaling factor (a.k.a., noise amplification factor) of the precoder [17]. Accordingly, VP achieves a significant gain in the error probability compared to ZF at high SNR regimes. To be more specific, the transmit signal vector with VP

precoding is formed as

$$\mathbf{x}_{VP} = \frac{1}{f_{VP}} \cdot \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} (\mathbf{s} + \tau \cdot \mathbf{l}), \tag{2.22}$$

where $\tau = 2|c|_{\max} + \Delta$ is the modulation-dependent modulo basis. $|c|_{\max}$ is equal to the absolute value of the constellation points with the maximum magnitude, and $\Delta$ is the minimum distance between constellation points, respectively. It is worth noting that the real part and imaginary part of the constellation points should be considered separately in calculating $\tau$ [17]. For example, for a normalised Quadrature PSK (QPSK) modulation, $\tau_{QPSK} = \frac{2\times 1+2}{\sqrt{2}} = 2.8284$, while for a normalised 16 Quadrature Amplitude Modulation (16QAM), $\tau_{16QAM} = \frac{2\times 3+2}{\sqrt{10}} = 2.5298$. $\mathbf{l} \in \mathbb{C}\mathbb{Z}^{K\times 1}$ denotes the complex-integer perturbation vector, which is obtained by the sphere encoding [17], given by

$$\mathbf{l} = \underset{\mathbf{l}\in\mathbb{C}\mathbb{Z}^{K\times 1}}{\arg\min} \left\| \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1} (\mathbf{s} + \tau \cdot \mathbf{l}) \right\|_F^2. \tag{2.23}$$

Based on (2.22), the scaling factor for VP is obtained as

$$f_{VP} = \sqrt{\mathrm{tr}\left\{ (\mathbf{s} + \tau \cdot \mathbf{l})^H (\mathbf{H}\mathbf{H}^H)^{-1} (\mathbf{s} + \tau \cdot \mathbf{l}) \right\}}. \tag{2.24}$$

With VP precoding, the received signal for user $k$ can be expressed as

$$\begin{aligned} y_k &= \frac{1}{f_{VP}} \cdot \mathbf{h}_k \mathbf{x}_{VP} + n_k \\ &= \frac{1}{f_{VP}} (s_k + \tau l_k) + n_k, \end{aligned} \tag{2.25}$$

where $\mathbf{l} = [l_1, l_2, \cdots, l_K]^T$. In order to remove the perturbation factor $\tau l_k$, a modulo operation must be performed for each user after the received signal is rescaled. Accordingly, the received signal of user $k$ for demodulation can

be expressed as

$$\begin{aligned} r_k &= \mod_\tau \{f_{VP} y_k\} \\ &= \mod_\tau \{s_k + \tau l_k + f_{VP} n_k\} \\ &= s_k + f_{VP} \hat{n}_k, \end{aligned} \tag{2.26}$$

where $\hat{n}_k$ represents the effective noise that considers the modulo loss effect.

As the modulo basis $\tau$ is modulation dependent, the standard VP precoding technique described above is only applicable to single-modulation scenarios. Modified VP approaches are proposed in [18, 48] to circumvent this drawback, such that VP precoding can be applied to multiple-modulation scenarios, where the data symbols are from different modulation constellations. Nevertheless, both techniques proposed in [18, 48] can only achieve a sub-optimal VP performance, which motivates our proposed techniques that achieve an improved error probability, included in Chapter 3 of this thesis.

### 2.2.3 Optimisation-based Precoding Techniques

In addition to the introduced linear and non-linear precoding methods, in the following we also briefly describe the precoding approaches, which are based on optimisations on the received SINR. For such optimisation-based precoding, the signal vector $\mathbf{x}$ is usually expressed as

$$\mathbf{x} = \sum_{i=1}^{K} \mathbf{w}_i s_i, \tag{2.27}$$

where each $\mathbf{w}_i$ denotes the precoding vector for the corresponding data symbol $s_i$ and unit power is assumed for each data symbol. Accordingly, (2.10) can be further expressed as

$$\begin{aligned} y_k &= \mathbf{h}_k \sum_{i=1}^{K} \mathbf{w}_i s_i + n_k \\ &= \mathbf{h}_k \mathbf{w}_k s_k + \mathbf{h}_k \sum_{i \neq k}^{K} \mathbf{w}_i s_i + n_k, \end{aligned} \tag{2.28}$$

based on which we express the received SINR of the $k$-th user as

$$\gamma_k = \frac{|\mathbf{h}_k \mathbf{w}_k|^2}{\sum\limits_{i \neq k}^{K} |\mathbf{h}_k \mathbf{w}_i|^2 + \sigma^2}. \tag{2.29}$$

**Power Minimisation** is a popular optimisation-based precoding method, where we minimise the required transmit power of the BS subject to a predefined SINR target for each user. This leads to the mathematical construction of the optimisation problem as

$$\begin{aligned} \mathscr{P}_{2.1}: \ & \min_{\mathbf{w}_i} \ \sum_{i=1}^{K} \|\mathbf{w}_i\|_F^2 \\ s.t. \ & \frac{|\mathbf{h}_k \mathbf{w}_k|^2}{\sum\limits_{i \neq k}^{K} |\mathbf{h}_k \mathbf{w}_i|^2 + \sigma^2} \geq \Gamma_k, \ \forall k \in \{1, 2, \cdots, K\} \end{aligned} \tag{2.30}$$

where each $\Gamma_k$ denotes the SINR target of user $k$. $\mathscr{P}_{2.1}$ is a convex optimisation problem and can be efficiently solved as a virtual uplink power allocation problem or via semidefinite optimisation methods [39].

**SINR Balancing** precoding targets at maximising the minimum received SINR for the users, subject to a predefined maximum available transmit power. Mathematically, this problem is formulated as

$$\begin{aligned} \mathscr{P}_{2.2}: \ & \max_{\mathbf{w}_i} \min_{k} \ \gamma_k \\ s.t. \ & \gamma_k = \frac{|\mathbf{h}_k \mathbf{w}_k|^2}{\sum\limits_{i \neq k}^{K} |\mathbf{h}_k \mathbf{w}_i|^2 + \sigma^2}, \ \forall k \in \{1, 2, \cdots, K\} \\ & \sum_{i=1}^{K} \|\mathbf{w}_i\|_F^2 \leq P_0 \end{aligned} \tag{2.31}$$

where $P_0$ denotes the maximum transmit power available at the BS. Different from $\mathscr{P}_{2.1}$, $\mathscr{P}_{2.2}$ is a non-convex optimisation. Fortunately, its optimal solution can be obtained either through a bisection search method described in [40], or via an iterative algorithm explained in [49].

### 2.2.4 Constructive Interference Precoding

The aforementioned precoding approaches operate at a frame level, i.e., the precoding matrices are independent of the data symbols. Moreover, these techniques are aimed at interference minimisation. In this section, we further introduce a symbol-level data-dependent downlink precoding technique that further improves the performance, based on the fact that instantaneous interference can be beneficial and further exploited on a symbol level [50, 51].

CI is defined as the interference that pushes the received signals away from the detection thresholds of the modulation constellations. Pioneering works in [52, 53] have shown that instantaneous interference can be classified into constructive and destructive. A closed-form linear precoding method that exploits the CI while eliminates the destructive part has been introduced in [54], and a more advanced linear approach is proposed in [55], where the phases of destructive interference are controlled and rotated such that the destructive interference becomes constructive. Based on this phase rotation concept, the interference exploitation is further extended to the areas of optimisation-based precoding [41], where the constructive region is introduced. To illustrate this intuitively, in Fig. 2.3 we depict the constructive region for several modulation types, where the green areas denote the constructive region. $\theta_t$ is an introduced variable for the construction of the optimisation problem, which will be described in (2.36) in the following. As can be observed, when the received signals fall within the constructive region, they are farther away from the detection thresholds, which promise an improved detection performance. On the other hand, when located in the destructive region, they are closer to the detection thresholds.

**The Phase-Rotation Formulation:** In the following, we construct the precoding optimisation based on CI for both power minimisation and SINR balancing. Without loss of generality, we consider a normalised QPSK constellation and focus on the constellation point $\left(\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{2}}\cdot j\right)$, as shown in

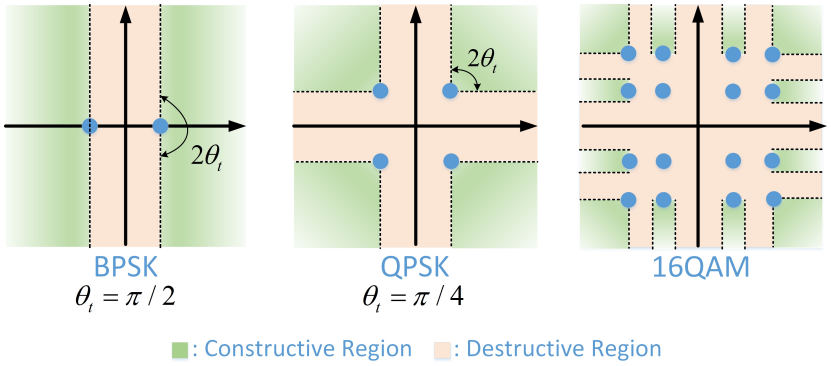



**Figure 2.3:** Examples of constructive region for different modulations

Fig. 2.4. Based on [41], for a CI-based power minimisation we denote

$$\vec{OA} = \sqrt{\Gamma_k \sigma^2} \cdot s_k, \tag{2.32}$$

and we denote $\vec{OB}$ as the noiseless received signal for user $k$, which is expressed as

$$\vec{OB} = \mathbf{h}_k \mathbf{W} \mathbf{s} = \lambda_k s_k, \tag{2.33}$$

where $\lambda_k$ is an introduced complex variable that represents the effect of multi-user interference on user $k$. If $\lambda_k$ is purely real, it means that multi-user interference is zero or strictly aligned to the data symbol of interest. Geometrically, we can further obtain the expression of $\vec{AC}$ and $\vec{BC}$, given by

$$\vec{AC} = \left[\Re(\lambda_k) - \sqrt{\Gamma_k \sigma^2}\right] s_k, \ \vec{BC} = j \cdot \Im(\lambda_k) s_k, \tag{2.34}$$

where the imaginary unit $j$ represents a phase rotation of 90° along an anti-clockwise direction.

Subsequently, that the node 'B' is located in the constructive region is

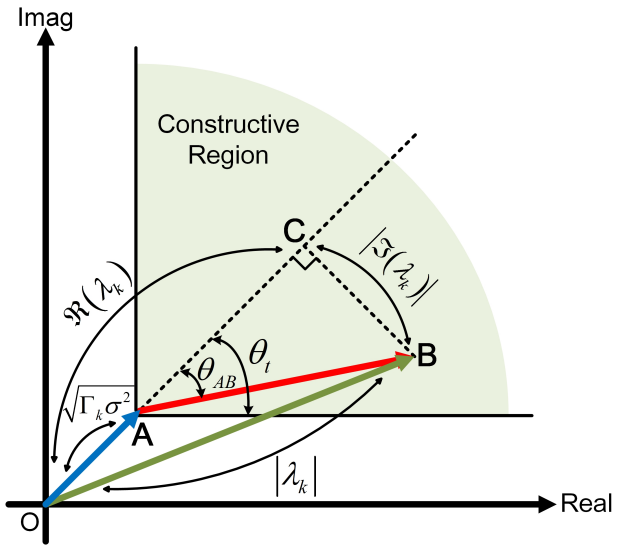



**Figure 2.4:** Constructive region, QPSK

equivalent to the following condition:

$$\begin{aligned} &\theta_{AB} \leq \theta_t \Rightarrow \tan\theta_{AB} \leq \tan\theta_t \\ \Rightarrow &\frac{|j\cdot\Im(\lambda_k)s_k|}{\left|\left[\Re(\lambda_k)-\sqrt{\Gamma_k\sigma^2}\right]s_k\right|} \leq \tan\theta_t \\ \Rightarrow &\left[\Re(\lambda_k)-\sqrt{\Gamma_k\sigma^2}\right]\tan\theta_t \geq |\Im(\lambda_k)|, \end{aligned} \tag{2.35}$$

where $\theta_t$ denotes the angle threshold and is geometrically obtained as

$$\theta_t = \frac{\pi}{\mathbb{N}} \tag{2.36}$$

for $\mathbb{N}$-PSK modulations, shown in both Fig. 2.3 and Fig. 2.4. Accordingly, we can formulate the power minimisation problem based on interference exploitation as

$$\begin{aligned} \mathscr{P}_{2.3}: \ &\min_{\mathbf{W}} \ \|\mathbf{Ws}\|_F^2 \\ s.t. \ \ &\mathbf{h}_k\mathbf{Ws} = \lambda_k s_k, \ \forall k \in \{1,2,\cdots,K\} \\ &\left[\Re(\lambda_k)-\sqrt{\Gamma_k\sigma^2}\right]\tan\theta_t \geq |\Im(\lambda_k)|, \ \forall k \in \{1,2,\cdots,K\} \end{aligned} \tag{2.37}$$

which is a convex optimisation. Following a similar step, the CI-based SINR

balancing optimisation can be formulated as

$$\begin{aligned}\mathscr{P}_{2.4}: \ & \max_{\mathbf{W}} \min_{k} \ \Gamma_k \\ s.t. \quad & \mathbf{h}_k \mathbf{W}\mathbf{s} = \lambda_k s_k, \ \forall k \in \{1,2,\cdots,K\} \\ & \left[\Re\left(\lambda_k\right) - \sqrt{\Gamma_k \sigma^2}\right] \tan\theta_t \geq \left|\Im\left(\lambda_k\right)\right|, \ \forall k \in \{1,2,\cdots,K\} \\ & \|\mathbf{W}\mathbf{s}\|_F^2 \leq P_0\end{aligned} \tag{2.38}$$

which can be further simplified into

$$\begin{aligned}\mathscr{P}_{2.5}: \ & \max_{\mathbf{W},t} \ t \\ s.t. \quad & \mathbf{h}_k \mathbf{W}\mathbf{s} = \lambda_k s_k, \ \forall k \in \{1,2,\cdots,K\} \\ & \left[\Re\left(\lambda_k\right) - t\right] \tan\theta_t \geq \left|\Im\left(\lambda_k\right)\right|, \ \forall k \in \{1,2,\cdots,K\} \\ & \|\mathbf{W}\mathbf{s}\|_F^2 \leq P_0\end{aligned} \tag{2.39}$$

Geometrically, $\mathscr{P}_{2.5}$ is interpreted as an optimisation where we maximise the distance between the constructive region and the detection thresholds such that the received signals are pushed as far away as possible.

Different from generic SINR balancing optimisation that is non-convex, CI-based SINR balancing optimisation is convex and can be efficiently solved. In both $\mathscr{P}_{2.3}$ and $\mathscr{P}_{2.5}$, we further note a key difference compared to the conventional frame-level precoding problems $\mathscr{P}_{2.1}$ and $\mathscr{P}_{2.2}$: the symbol-level precoding matrix $\mathbf{W}$ is dependent on the data symbol vector $\mathbf{s}$, and therefore the transmit power should be expressed as $\|\mathbf{W}\mathbf{s}\|_F^2$ and optimised on a symbol level, which is different from $\mathscr{P}_{2.1}$ and $\mathscr{P}_{2.2}$ where an average transmit power is maintained.

**The Symbol-Scaling Formulation:** The above CI formulations focus on PSK modulations only, while the extension to QAM has been included in [50, 51]. Based on Fig. 2.3, the CI can only be exploited by the outer constellation points for QAM modulations, while all the interference is destructive for the inner constellation points. Accordingly, in the following we reformulate the

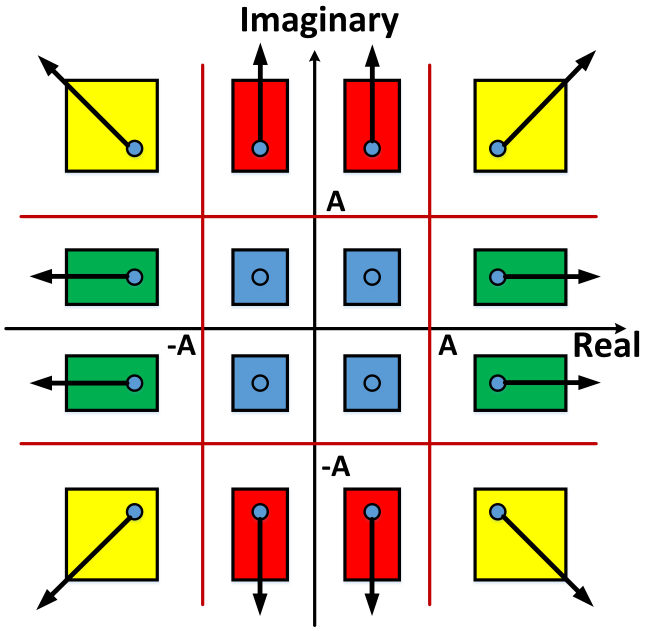



**Figure 2.5:** Constructive region, 16QAM

CI constraints based on a symbol-scaling formulation, which is applicable for QAM modulations. We decompose the noiseless received signal for user $k$ into

$$\mathbf{h}_k\mathbf{W}\mathbf{s} = \lambda_k^{\Re}\Re\left(s_k\right) + j\cdot\lambda_k^{\Re}\Im\left(s_k\right),\ \forall k \in \{1,2,\cdots,K\}, \tag{2.40}$$

where $\lambda_k^{\Re} \geq 0$ and $\lambda_k^{\Im} \geq 0$ are real scaling factors for the real and imaginary part of $s_k$, respectively. By expanding the received signal vector with its real and imaginary part, mathematically we obtain

$$\begin{bmatrix} \Re\left(\mathbf{HWs}\right) \\ \Im\left(\mathbf{HWs}\right) \end{bmatrix} = \mathbf{\Phi}\begin{bmatrix} \Re\left(\mathbf{s}\right) \\ \Im\left(\mathbf{s}\right) \end{bmatrix}, \tag{2.41}$$

where $\mathbf{\Phi} = diag\left\{\left[\lambda_1^{\Re},\cdots,\lambda_K^{\Re},\lambda_1^{\Im},\cdots,\lambda_K^{\Im}\right]\right\}$ represents a diagonal symbol-scaling matrix. In the following, we employ a normalised 16QAM constellation as an example, as shown in Fig. 2.5. $A = 0.6324$ denotes the detection threshold for this constellation. Accordingly, we categorise the constellation points for 16QAM into 4 types:

- Yellow: $\Re\left(s_k\right) \geq A$ and $\Im\left(s_k\right) \geq A$. This case is similar to the CI formu-

lation for PSK modulations, and we obtain $\lambda_k^{\Re} \geq 0$ and $\lambda_k^{\Im} \geq 0$;

- Red: $\Re(s_k) \leq A$ and $\Im(s_k) \geq A$. In this case, the interference can only be exploited for the imaginary part, and we obtain $\lambda_k^{\Re} = 1$ and $\lambda_k^{\Im} \geq 0$;
- Green: $\Re(s_k) \geq A$ and $\Im(s_k) \leq A$. In this case, CI can only be exploited for the real part, and we obtain $\lambda_k^{\Re} \geq 0$ and $\lambda_k^{\Im} = 1$;
- Blue: $\Re(s_k) \leq A$ and $\Im(s_k) \leq A$. All interference is destructive in this case, and we obtain $\lambda_k^{\Re} = 1$ and $\lambda_k^{\Im} = 1$.

By incorporating the above CI constraints, the optimisation problem aimed at power minimisation or SINR balancing for QAM modulations can be similarly constructed, and is omitted for brevity.

In 2.1, we summarise the key differences for the precoding techniques introduced in this section.

## 2.3 Channel Modelling

It can be seen in Chapter 2.2 that the knowledge of the channel is essential for all precoding techniques, and therefore the channel modelling is crucial in the analyses of the precoding techniques. In this section, we briefly introduce the channel models that are relevant to the contributions of this thesis, where throughout this thesis we consider the non-line-of-sight (NLoS) prop-

| Linear Precoding | | | |
|---|---|---|---|
| Name | Closed-form Expression | Complexity | Performance |
| MF | $\mathbf{P}_{MF} = \frac{1}{f_{MF}} \cdot \mathbf{H}^H$ | Simplest | Lowest |
| ZF | $\mathbf{P}_{ZF} = \frac{1}{f_{ZF}} \cdot \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H\right)^{-1}$ | Higher than MF | Better than MF |
| RZF | $\mathbf{P}_{RZF} = \frac{1}{f_{RZF}} \cdot \mathbf{H}^H \left(\mathbf{H}\mathbf{H}^H + \alpha \cdot \mathbf{I}\right)^{-1}$ | Similar to ZF | Better than ZF |
| Non-linear Precoding | | | |
| Name | Expression | Complexity | Processing at Receiver? |
| THP | (2.16) - (2.21) | Higher than linear | Yes |
| VP | (2.22) - (2.24) | Higher than linear | Yes |
| Optimisation-based Precoding | | | |
| Optimisation Target | Problem Formulation | Convexity? | Solution |
| Power Minimisation | $\mathscr{P}_{2.1}$ | Yes | Standard |
| SINR Balancing | $\mathscr{P}_{2.2}$ | No | Bisection search |
| Interference Exploitation | $\mathscr{P}_{2.3}$, $\mathscr{P}_{2.4}$ | Yes | Standard |

**Table 2.1:** Classification of transmit processing techniques in the downlink

agation channels with flat fading, unless otherwise stated.

### 2.3.1 Uncorrelated Rayleigh Channel

The uncorrelated Rayleigh channel is the most common and widely adopted channel model. This model assumes the existence of large-scale statistically independent reflectors and scatters in the wireless environment, based on which each path of the channel tap can be modelled as a complex random variable. Subsequently, each channel tap is the summation of many independent complex random variables, which therefore follows a complex Gaussian distribution according to the Central Limit Theorem [3]. Without loss of generality, each channel coefficient can be modelled as

$$[\mathbf{H}]_{m,n} \sim \mathbb{CN}(0,1). \tag{2.42}$$

Based on the above modelling, the magnitude of each channel tap follows a Rayleigh distribution, which names this channel model.

### 2.3.2 Correlated Rayleigh Channel

The above uncorrelated channel model assumes that the channels between different antennas are statistically independent and ignores the spatial correlation effect at the antenna array. In fact for a practical antenna array, the channels between adjacent antennas are correlated, especially when the antenna spacing is small compared to the carrier wavelength [56, 57].

**Separately-Correlated Rayleigh Channel:** When a P2P MIMO model is considered, there exists spatial correlation at both the transmitter and receiver. Following [58, 59, 60], we model the channel $\mathbf{H}$ as

$$\mathbf{H} = \mathbf{A}_r \mathbf{H}_\alpha \mathbf{A}_t^H. \tag{2.43}$$

In (2.43), $\mathbf{H}_\alpha$ is a diagonal matrix with complex Gaussian distributed entries that model the Rayleigh components of the channel, given by

$$\mathbf{H}_\alpha = \frac{1}{\sqrt{N_p}} \cdot diag\left(\alpha_1, \alpha_2, ..., \alpha_{N_p}\right), \tag{2.44}$$

where $N_p$ is the number of independent paths (scattering) in the wireless environment, and each $\alpha_m \sim \mathbb{CN}(0,1)$, $m \in \left\{1, 2, \cdots, N_p\right\}$. $\mathbf{A}_r \in \mathbb{C}^{N_r \times N_p}$ and $\mathbf{A}_t \in \mathbb{C}^{N_t \times N_p}$ denote the receive-side and transmit-side steering matrices that contain $N_p$ steering vectors of the antenna array to model the spatial correlation effect. $\mathbf{A}_r$ and $\mathbf{A}_t$ can be decomposed into

$$\begin{aligned} \mathbf{A}_r &= \left[\mathbf{a}_r\left(\phi_1^r\right), \mathbf{a}_r\left(\phi_2^r\right), \cdots, \mathbf{a}_r\left(\phi_{N_p}^r\right)\right], \\ \mathbf{A}_t &= \left[\mathbf{a}_t\left(\phi_1^t\right), \mathbf{a}_t\left(\phi_2^t\right), \cdots, \mathbf{a}_t\left(\phi_{N_p}^t\right)\right], \end{aligned} \tag{2.45}$$

where for uniform linear arrays (ULAs), $\mathbf{a}_r\left(\phi_m^r\right) \in \mathbb{C}^{N_r \times 1}$ and $\mathbf{a}_t\left(\phi_m^t\right) \in \mathbb{C}^{N_t \times 1}$ can be expressed as [60]

$$\begin{aligned} \mathbf{a}_r\left(\phi_m^r\right) &= \left[1, e^{j2\pi d_r \sin\phi_m^r}, \cdots, e^{j2\pi(N_r-1)d_r \sin\phi_m^r}\right]^T, \; m \in \left\{1, 2, \cdots, N_p\right\}, \\ \mathbf{a}_t\left(\phi_m^t\right) &= \left[1, e^{j2\pi d_t \sin\phi_m^t}, \cdots, e^{j2\pi(N_t-1)d_t \sin\phi_m^t}\right]^T, \; m \in \left\{1, 2, \cdots, N_p\right\}. \end{aligned} \tag{2.46}$$

In (2.46), $\phi_m^r$ and $\phi_m^t$, $m \in \left\{1, 2, \cdots, N_p\right\}$ denote the angles of arrival (AoAs) and angles of departure (AoDs) respectively, and both $\phi_m^r$ and $\phi_m^t$ follow a Laplacian distribution [60]. $d_r$ and $d_t$ denote the equidistant spacing between adjacent antennas for the receive and transmit antenna array respectively, which is normalised by the carrier wavelength.

**Semi-Correlated Rayleigh Channel:** When a multi-user MIMO system is considered, it is reasonable that the considered channel is semi-correlated, where the spatial correlation only exists at the transmit side. In such case, each channel vector for the users is modelled as [61, 62]

$$\mathbf{h}_k = \mathbf{g}_k \mathbf{A}_k, \; \forall k \in \{1, 2, \cdots, K\}, \tag{2.47}$$

where $\mathbf{g}_k \in \mathbb{C}^{1\times N_s}$ and each entry of $\mathbf{g}_k$ follows the standard complex Gaussian distribution, which models the Rayleigh components. Each $\mathbf{A}_k \in \mathbb{C}^{N_s\times N_t}$ contains $N_s$ steering vectors that model the transmit correlation, where $N_s$ denotes the number of directions of departure (DoDs). For ULAs, each $\mathbf{A}_k$ can be modelled as

$$\mathbf{A}_k = \frac{1}{\sqrt{N_s}} \cdot \left[\mathbf{a}^T\left(\theta_{k,1}\right), \mathbf{a}^T\left(\theta_{k,2}\right), \cdots, \mathbf{a}^T\left(\theta_{k,N_s}\right)\right]^T, \ \forall k \in \{1,2,\cdots,K\}, \quad (2.48)$$

where each $\mathbf{a}^T\left(\theta_{k,i}\right) \in \mathbb{C}^{1\times N_t}$ is given by

$$\mathbf{a}\left(\theta_{k,i}\right) = \left[1, e^{j2\pi d\sin\theta_{k,i}}, \cdots, e^{j2\pi(N_t-1)d\sin\theta_{k,i}}\right]. \quad (2.49)$$

In (2.49), $d$ denotes the equidistant antenna spacing normalised by the carrier wavelength. Following the modelling for separately-correlated channels, each $\theta_{k,i}$ is assumed to follow a Laplacian distribution [60].

**Mutual Coupling Effect:** MC defines the electromagnetic interaction between the antenna elements in an antenna array [63]. When the spacing between adjacent antenna elements becomes smaller, the spatial correlation is not the only effect that has an impact on the performance of multiple-antenna systems. The coupling effect also arises as a key factor on which the system performance is dependent, based on the fact that the current on an antenna element is jointly decided by its own excitation as well as the contributions from other nearby antenna elements. In a multiple-antenna communication system, it has been shown in [61, 62, 64] that the MC effect can be fully characterised by a MC matrix $\mathbf{Z}$ that is multiplied to the channel matrix, whose formulation is derived in Appendix A.

### 2.3.3 Millimeter-Wave Channel

In addition to the above channel models, we also present the mmWave channel model in this section. Compared to aforementioned channels, mmWave channels are expected to have limited scattering [65, 66]. To take this effect into consideration, a geometric channel model with $L_u$ independent propa-

gation paths is considered for each $\mathbf{H}_k$, modelled as [67]

$$\mathbf{H}_k = \sqrt{\frac{N_t N_r^k}{L_u}} \sum_{l=1}^{L_u} \alpha_{k,l} \cdot \mathbf{a}_{MS}\left(\theta_{k,l}^{MS}, \phi_{k,l}^{MS}\right) \mathbf{a}_{BS}^{H}\left(\theta_{k,l}^{BS}, \phi_{k,l}^{BS}\right), \tag{2.50}$$

where the value of $L_u$ is small compared to $N_t$ for limited scattering. $\alpha_{k,l}$ is the complex path gain, where $\alpha_{k,l}$ represents a deterministic complex channel gain for the LoS components, while $\alpha_{k,l}$ follows the standard complex Gaussian distribution for the NLoS components. $\mathbf{a}_{MS}\left(\theta_{k,l}^{MS}, \phi_{k,l}^{MS}\right)$ and $\mathbf{a}_{BS}^{H}\left(\theta_{k,l}^{BS}, \phi_{k,l}^{BS}\right)$ denote the receive and transmit array response vectors respectively, where $\theta_{k,l}^{MS}$ ($\phi_{k,l}^{MS}$) and $\theta_{k,l}^{BS}$ ($\phi_{k,l}^{BS}$) denote the azimuth (elevation) angles of arrival and departure (AoAs/AoDs) of the $l$-th path. For uniform linear arrays (ULAs), only the azimuth AoAs and AoDs are considered, and we can express the array response as

$$\mathbf{a}_{ULA}(\theta) = \frac{1}{\sqrt{N_t}}\left[1, e^{j\cdot kd\sin(\theta)}, \cdots, e^{j\cdot(N_t-1)kd\sin(\theta)}\right]^T, \tag{2.51}$$

where $k = \frac{2\pi}{\lambda}$ with $\lambda$ being the carrier wavelength, and $d$ is the physical inter-antenna spacing. When uniform planar arrays (UPAs) are considered, the array response vector can be expressed as

$$\begin{aligned}\mathbf{a}_{UPA}(\theta,\phi) = \frac{1}{\sqrt{N_{tot}}}\Big[1, \cdots, e^{j\cdot kd[m\sin(\phi)\sin(\theta)+n\cos(\theta)]}, \\ \cdots, e^{j\cdot kd[(W-1)\sin(\phi)\sin(\theta)+(H-1)\cos(\theta)]}\Big]^T,\end{aligned} \tag{2.52}$$

where $0 \le m \le W-1$, $0 \le n \le H-1$, and $N_{tot} = WH$. $N_{tot} = N_t$ for $\mathbf{a}_{BS}$ and $N_{tot} = N_r^k$ for $\mathbf{a}_{MS}$ of $\mathbf{H}_k$.

### 2.3.4 Modelling of Imperfect CSI

It is observed in Chapter 2.3 that downlink precoding techniques require the knowledge of the channel at the BS. In practical cellular communication systems, the perfect acquisition of the channel knowledge is difficult, and it is essential to consider the performance of downlink transmission methods

under imperfect CSI. Depending on the duplex mode, we present the models of imperfect CSI employed in the following chapters of this thesis.

**Statistical CSI Error Model:** In the time division duplex (TDD) mode, the uplink and downlink channels operate in the same frequency bands, and therefore the downlink channel can be directly measured by the uplink-downlink reciprocity at the BS, which is subject to estimation errors at the BS [68]. When the channel estimation technique based on minimum mean squared error (MMSE) is employed, as described in [69, 70], the imperfect channel model is expressed as

$$\mathbf{H} = \alpha \cdot \left(\hat{\mathbf{H}} + \mathbf{E}\right) + \mathbf{Q}, \tag{2.53}$$

where $\mathbf{H}$ represents the real wireless propagation channel, and $\hat{\mathbf{H}}$ is the estimated channel at the BS. $\alpha$ denotes a correlation factor that models the time delay in estimating the channel. $\alpha = 1$ if the channel that the precoded signals experience is identical to the estimated channel based on which the signals are precoded. $\mathbf{Q}$ denotes the corresponding delay error matrix, where each entry in $\mathbf{Q}$ is independent and distributed as $\mathbb{CN}\left(0, 1-\alpha^2\right)$. $\mathbf{E}$ represents the estimation error matrix with $\mathbf{E} \sim \mathbb{CN}(0, \eta \cdot \mathbf{I})$, and $\eta$ is given by

$$\eta = \frac{\beta}{\rho}, \tag{2.54}$$

where $\rho$ denotes the transmit SNR and $\beta$ is the inverse proportionality coefficient. When $\alpha = 1$, we arrive at a simplified imperfect CSI model for TDD systems as

$$\mathbf{H} = \hat{\mathbf{H}} + \mathbf{E}, \tag{2.55}$$

which is employed in Chapter 6. The perfect CSI case is equivalent to $\alpha = 1$ and $\beta = 0$.

**Norm-Bounded CSI Error Model:** In frequency division duplex (FDD) mode, as the downlink channel and uplink channel are operating in different frequency bands, the uplink-downlink reciprocity does not hold. In practi-

cal systems, the channel knowledge is first estimated at the receivers, and then fedback to the BS [68]. Accordingly, the CSI errors for FDD systems are dominated by the quantisation errors in the limited feedback, based on which the imperfect CSI model can be constructed as

$$\mathbf{h}_k = \hat{\mathbf{h}}_k + \mathbf{e}_k,\ \forall k \in \{1,2,\cdots,K\}, \tag{2.56}$$

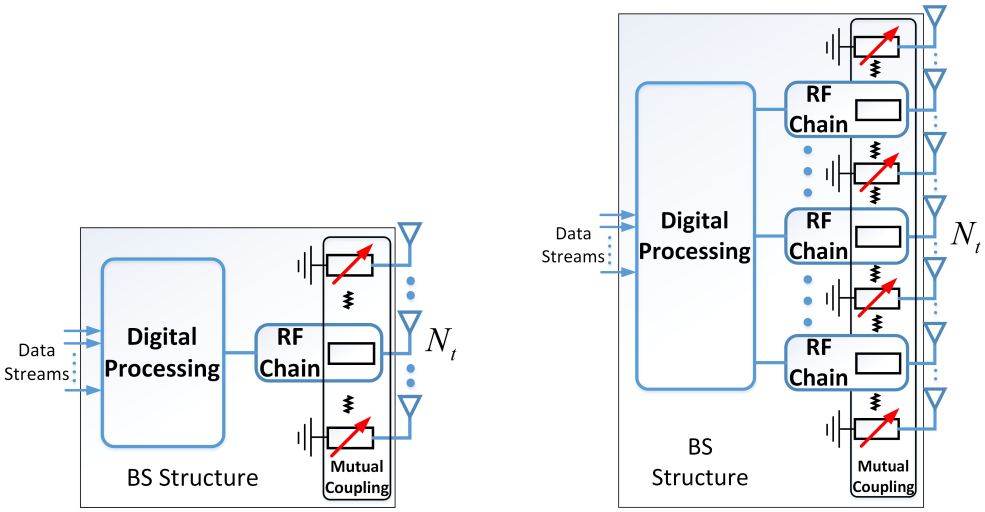
where for each user $k$, due to the quantisation the channel uncertainty is bounded by a spherical region, expressed as

$$\mathscr{D}_k := \left\{ \hat{\mathbf{h}}_k + \mathbf{e}_k \mid \|\mathbf{e}_k\|_F \leq \sqrt{\delta_k} \right\},\ \forall k \in \{1,2,\cdots,K\}. \tag{2.57}$$

## 2.4 Hybrid Multiple-Antenna Structures

The precoding techniques introduced in Chapter 2.3 require a FD structure at the BS, i.e., both the phases and the amplitudes of the precoding matrices can be controlled and tuned arbitrarily [47]. Nevertheless, such FD systems may not be practical in the future 5G-and-beyond communication systems, which target at energy efficient transmissions. This is because the FD systems require the number of RF chains to be equal to the number of transmit antennas, which incur a significant increase in the power consumption, hardware complexity and hardware cost at the BS, especially for the large-scale multiple-antenna systems. Towards the future energy efficient communication systems, alternative hybrid BS structures have been considered, mainly realised by employing a limited number of RF chains.

In the following, we introduce two typical hybrid BS structures with a reduced number of RF chains, both of which are studied within this thesis. These hybrid structures either respect the existence of the MC effect for compact antenna arrays [8], or employ an analog processing network of phase shifters for reduced hardware complexity [23].




**(a)** Single-fed ESPAR **(b)** Multi-fed ESPAR

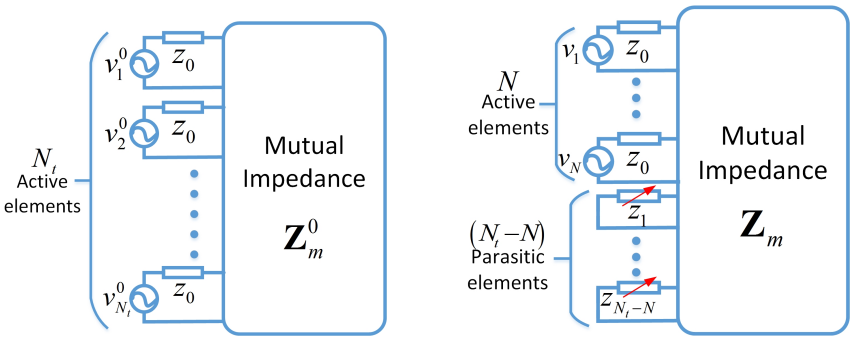
**Figure 2.6:** Example of a single-fed and multi-fed parasitic-antenna BS

### 2.4.1 Compact Parasitic-Antenna Systems

When a compact antenna array is considered, i.e., the equidistant antenna spacing is small, the strong MC effect usually has a negative effect on the system performance [63, 71, 72, 73]. Subsequently, extensive studies have been conducted to compensate for the performance losses due to the coupling effect [7, 20, 74, 75, 76, 77].

In addition, a novel parasitic-antenna array known as ESPAR has been proposed in [8, 14, 15], where the ESPAR-based arrays consist of limited active antenna elements with their own fixed antenna loads and RF chains, surrounded by parasitic elements with tunable antenna loads, which are excited passively. In Fig. 2.6, we present both a single-fed and multi-fed ESPAR-based BS structure, where it is observed that the number of RF chains is equal to that of active elements. Accordingly, the required number of RF chains is significantly reduced, which makes the parasitic-antenna arrays a space- and energy-efficient alternative to generic antenna array architectures.

Moreover, as opposed to the conventional multiple-antenna arrays where the small antenna spacing and the resulting coupling effect are detri-




**(a)** Conventional Antenna Array **(b)** Parasitic Antenna Array

**Figure 2.7:** Circuit representation of the conventional and parasitic antenna array

mental, the small size constraint turns to be an advantage for ESPAR-based MIMO systems. In particular, the parasitic arrays exploit the MC effect instead of compensating for this effect, where the desired radiation patterns are formed by controlling the voltages on the active elements and the loads on the parasitic elements. To be specific, we first introduce the signal model for the conventional antenna array, and further extend to the parasitic-antenna array.

**Conventional Antenna Array:** The circuit representation of a conventional FD antenna array is shown in Fig. 2.7 (a). In this case, each antenna element $k$ is fed by an independent source with the complex voltage $v_k$, and the fixed output impedance is denoted as $z_0$. The constant mutual impedance matrix for the conventional array is denoted as $\mathbf{Z}_m^0 \in \mathbb{C}^{N_t \times N_t}$, which is related to both the carrier frequency and the antenna spacing. According to the generalised Ohm's law, the complex current vector at the antenna port can be expressed as [78]

$$\mathbf{i}_0 = \left[ diag\left(\mathbf{z}_0\right) + \mathbf{Z}_m^0 \right]^{-1} \mathbf{v}_0, \tag{2.58}$$

where $\mathbf{i}_0 \in \mathbb{C}^{N_t \times 1}$ is the current vector, $\mathbf{z}_0 = z_0 \cdot \mathbf{1} \in \mathbb{C}^{N_t \times 1}$ is the fixed output impedance vector, and $\mathbf{v}_0 \in \mathbb{C}^{N_t \times 1}$ denotes the voltage vector. For a generic FD MIMO array, $z_0$ usually equals $50\Omega$, and the desired signal vector is generated by adjusting the voltage vector $\mathbf{v}_0$.

**Parasitic Antenna Array:** The equivalent circuit representation of a par-

asitic array is shown in Fig. 2.7 (b). Without loss of generality, we assume that the total number of active antennas is $N$ (which is equivalent to the number of RF chains), while the remaining $(N_t - N)$ elements are parasitic and excited passively by the MC effect between antenna elements. In this case, the current vector at the antenna port can be obtained based on (2.58) as [79]

$$\begin{aligned}\mathbf{i} &= [diag(\mathbf{z}_L) + \mathbf{Z}_m]^{-1} \cdot [v_1, \cdots, v_N, 0, \cdots, 0]^T \\ &= [diag(\mathbf{z}_L) + \mathbf{Z}_m]^{-1} \mathbf{v}_s \\ &= \mathbf{Z}_T^{-1} \mathbf{v}_s,\end{aligned} \tag{2.59}$$

where we denote $\mathbf{Z}_T = [diag(\mathbf{z}_L) + \mathbf{Z}_m]$ as the effective coupling matrix. The antenna load vector $\mathbf{z}_L$ is expressed as

$$\mathbf{z}_L = [z_1, z_2, \cdots z_{N_t}]^T, \ z_n = z_0, \ \forall n \in \{1, 2, \cdots, N\}. \tag{2.60}$$

$z_0$ denotes the fixed load impedance that corresponds to the active antenna element, and each $z_i$, $\forall i \in \{N+1, N+2, \cdots, N_t\}$ denotes the tunable antenna load for each parasitic element. As observed from (2.59), different from the conventional antenna array where the transmit signals are solely controlled by the feeding voltages, the current vector for ESPARs is jointly decided by the voltages on the active antennas and the tunable loads on the parasitic antenna elements. By expanding (2.59), we obtain

$$\begin{bmatrix} [\mathbf{Z}_m]_{1,1} + z_1 & [\mathbf{Z}_m]_{1,2} & \cdots & [\mathbf{Z}_m]_{1,N_t} \\ [\mathbf{Z}_m]_{2,1} & [\mathbf{Z}_m]_{2,2} + z_2 & \ddots & [\mathbf{Z}_m]_{2,N_t} \\ \vdots & \ddots & \ddots & \vdots \\ [\mathbf{Z}_m]_{N_t,1} & [\mathbf{Z}_m]_{N_t,2} & \cdots & [\mathbf{Z}_m]_{N_t,N_t} + z_{N_t} \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_{N_t} \end{bmatrix} = \begin{bmatrix} v_1 \\ \vdots \\ v_N \\ \mathbf{0}^{(N_t - N) \times 1} \end{bmatrix}, \tag{2.61}$$

based on which we can calculate the feeding voltage and the values of the tunable antenna loads as a function of the desired current vector, given by

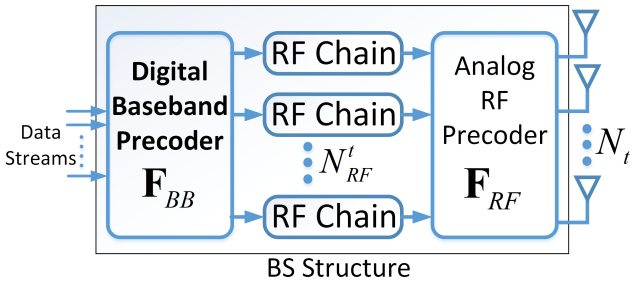



**Figure 2.8:** A generic hybrid AD BS structure

[8, 78]

$$\begin{aligned} v_n &= \sum_{k=1}^{N_t} [\mathbf{Z}_m]_{n,k} i_k + z_0 i_n, \ \forall n \in \{1,2,\cdots,N\}, \\ z_m &= -\frac{1}{i_m} \sum_{k=1}^{N_t} [\mathbf{Z}_m]_{m,k} i_k, \ \forall m \in \{N+1, N+2, \cdots, N_t\}. \end{aligned} \tag{2.62}$$

The precoding design for parasitic-antenna arrays will be studied in Chapter 6.

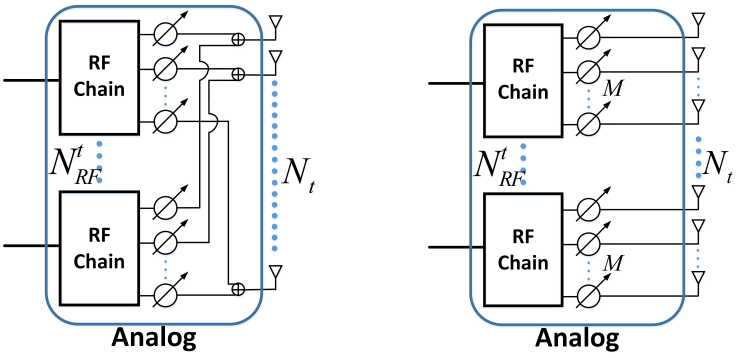
### 2.4.2 Hybrid Analog-Digital Structures

In addition to the parasitic-antenna arrays, we introduce another popular hybrid structure, which is initially designed for large-scale mmWave communications to alleviate the hardware cost [11, 23]. As depicted in Fig. 2.8, the BS employs a reduced number of $N_{RF}^t$ RF chains compared to the total number of $N_t$ transmit antennas. Accordingly, the downlink processing has been divided into the analog domain and the digital domain, where the precoding $\mathbf{F}$ consists of two parts: the high-dimensional analog precoding $\mathbf{F}_{RF}$, and the low-dimensional digital precoding $\mathbf{F}_{BB}$, expressed as

$$\mathbf{F} = \mathbf{F}_{RF}\mathbf{F}_{BB}, \tag{2.63}$$

where $\mathbf{F}_{RF} \in \mathbb{C}^{N_t \times N_{RF}^t}$ and $\mathbf{F}_{BB} \in \mathbb{C}^{N_{RF}^t \times K}$.

In the analog domain, the use of an analog network of phase shifters is the most popular one [11, 23, 24, 67, 80], which enforces a constant modulus constraint on the entries of the analog precoder. In addition, analog precoding based on switches is also a potential way to implement the analog




**(a)** Fully-Connected **(b)** Partially-Connected

**Figure 2.9:** Transmit Antenna Array Structure for Hybrid AD Precoding

processing [81]. Recently, another hybrid precoding based on lens antenna arrays has further been introduced as a cost-effective solution to large-scale multiple-antenna systems [82, 83]. In the digital domain, a low-dimensional conventional FD precoder is employed to manage the multiuser interference.

In this thesis, we focus on the hybrid design where the analog precoding is implemented via phase shifters. Two typical analog precoding structures, i.e., fully-connected and partially-connected structures are considered, as depicted in Fig. 2.9. When a fully-connected analog structure is considered, as shown in Fig. 2.9 (a), $\mathbf{F}_{RF}$ can be expressed as

$$\mathbf{F}_{RF} = \left[\mathbf{f}_1^f, \mathbf{f}_2^f, \cdots, \mathbf{f}_K^f\right], \tag{2.64}$$

which is a full matrix with all elements being non-zero and normalised to satisfy

$$\left|\left[\mathbf{f}_k^f\right]_m\right| = \frac{1}{\sqrt{N_t}}, \ \forall k \in \{1,2,\cdots,K\}, \ \forall m \in \{1,2,\cdots,N_t\}. \tag{2.65}$$

For partially-connected structures, as shown in Fig. 2.9 (b), $\mathbf{F}_{RF}$ can be decomposed into

$$\mathbf{F}_{RF} = diag\left\{\left[\mathbf{f}_1^p, \mathbf{f}_2^p, \cdots, \mathbf{f}_K^p\right]\right\}, \tag{2.66}$$

which is a block-diagonal matrix with each $\mathbf{f}_k^p \in \mathbb{C}^{M_p \times 1}$, where $M_p$ is the number of antennas connected to each RF chain. Each entry in $\mathbf{f}_k^p$ is normalised

to satisfy

$$\left|\left[\mathbf{f}_k^p\right]_m\right| = \frac{1}{\sqrt{M_p}},\ \forall k \in \{1,2,\cdots,K\},\ \forall m \in \left\{1,2,\cdots,M_p\right\}. \tag{2.67}$$

The specific design of hybrid precoding and its extension to energy efficient MIMO SWIPT are discussed in Chapter 7.

# Chapter 3

# VP Precoding for Multiple-Modulation Scenarios

This chapter is based on our works published in [J7], [J8].

## 3.1 Joint Vector Perturbation

### 3.1.1 Introduction

In this chapter, we focus on the study of non-linear VP techniques. Compared to linear precoding approaches, the non-linear VP precoding can achieve a greatly improved BER performance, and has received extensive research attention [17, 18, 48]. Nevertheless, as mentioned in Chapter 2 that the modulo basis $\tau$ is modulation dependent, one main drawback of the conventional VP method in [17] is its inapplicability in multiple-modulation scenarios where different users employ different modulation types.

To overcome this drawback, in [48] the conventional VP precoding proposed in [17] is combined with the block diagonalisation (BD) scheme designed for multiuser MIMO (MU-MIMO) systems, and a block-diagonalised VP (BDVP) approach is proposed, where the sphere search is performed for each user independently to obtain the perturbation vectors. A low-complexity BDVP approach is introduced in [18], where a user grouping VP (UGVP) precoding is further proposed to improve the performance of BDVP. For UGVP, the sphere search is performed for each modulation inde-

pendently instead of each user. While both BDVP and UGVP enable the employment of VP approaches in multiple-modulation scenarios, the above two methods can only achieve sub-optimal performances, as the sphere search algorithm is not jointly conducted for all the data streams.

In order to retain the performance advantages of the conventional VP precoding for single modulation, in this section we propose a JVP precoding technique that achieves a comparable performance to the conventional VP approach, which is applicable to multiple-modulation scenarios. With the introduced constellation scaling approach, different modulation types can employ an identical modulo basis, and therefore the sphere search for the perturbation vectors can be jointly obtained as in the conventional VP formulation. In addition to the performance superiority, another advantage of the proposed technique is that it forms a basis for other VP-based techniques to be applicable in multiple-modulation cases. For example, complexity-reduction methods proposed in [35, 86] can be applied upon the proposed JVP precoding to achieve further complexity benefits.

### 3.1.2 System Model

We begin with the description of the system model, followed by a brief introduction of the existing BDVP and UGVP approaches.

#### 3.1.2.1 MU-MIMO System Model

We consider a MU-MIMO downlink system, where a BS with $N_t$ transmit antennas is communicating with $K$ users simultaneously. Each user $k$ is equipped with $n_k^r \geq 1$ receive antennas, and we assume the number of streams for user $k$ is equal to $n_k^r$. The total number of receive antennas is obtained as $N_r = \sum_{k=1}^{K} n_k^r$ and $N_r \leq N_t$. We consider an uncorrelated Rayleigh flat-fading channel as introduced in Chapter 2.3.1, where the channel matrix for user $k$ is denoted as $\mathbf{H}_k \in \mathbb{C}^{n_k^r \times N_t}$, and the channel matrix between the BS and all users can be expressed in a compact form as $\mathbf{H} = \left[\mathbf{H}_1^T, \mathbf{H}_2^T, ..., \mathbf{H}_K^T\right]^T$.

Accordingly, the received signal vector for user $k$ can be expressed as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{F}_k \mathbf{x}_k + \mathbf{H}_k \sum_{i=1, i\neq k}^{K} \mathbf{F}_i \mathbf{x}_i + \mathbf{n}_k, \tag{3.1}$$

where $\mathbf{x}_k$ is the perturbed signal vector for user $k$, and $\mathbf{n}_k$ is the additive complex Gaussian noise vector at user $k$ with zero mean and covariance $\sigma^2 \cdot \mathbf{I}$. In (3.1), $\mathbf{F}_k$ is the precoding matrix for user $k$.

In multiple-modulation scenarios, the conventional VP approach is not directly applicable, as the search for the perturbation vectors for different modulation types cannot be jointly performed, which is due to the different values of the modulo basis $\tau$. In the following, we briefly review two existing VP-based approaches for multiple-modulation scenarios.

#### 3.1.2.2 BDVP Precoding

For BDVP introduced in [48], the precoding matrix $\mathbf{F}_k$ in (3.1) for the $k$-th user can be expressed as $\mathbf{F}_k = \mathbf{B}_k \mathbf{D}_k$, where $\mathbf{B}_k$ is the matrix used to eliminate the inter-user interference, which satisfies

$$\mathbf{H}_i \mathbf{B}_k = \mathbf{0}, \ \forall i \neq k. \tag{3.2}$$

$\mathbf{B}_k$ can be obtained from the SVD of $\tilde{\mathbf{H}}_k$ as

$$\tilde{\mathbf{H}}_k = \tilde{\mathbf{U}}_k \tilde{\mathbf{\Delta}}_k [\tilde{\mathbf{V}}_k^{(1)}, \tilde{\mathbf{V}}_k^{(0)}]^H, \tag{3.3}$$

where $\tilde{\mathbf{H}}_k = [\mathbf{H}_1^T, ..., \mathbf{H}_{k-1}^T, \mathbf{H}_{k+1}^T, ..., \mathbf{H}_K^T]^T$. Subsequently, $\mathbf{B}_k$ is chosen as

$$\mathbf{B}_k = \tilde{\mathbf{V}}_k^{(0)}, \tag{3.4}$$

as $\tilde{\mathbf{V}}_k^{(0)}$ forms the orthogonal basis of the null space of $\tilde{\mathbf{H}}_k$. After $\mathbf{B}_k$ is obtained, the equivalent channel for user $k$ is a non-interfering channel, which is expressed as

$$\mathbf{H}_k^{eq} = \mathbf{H}_k \mathbf{B}_k. \tag{3.5}$$

Accordingly, $\mathbf{D}_k$ is obtained based on ZF as

$$\mathbf{D}_k = (\mathbf{H}_k \mathbf{B}_k)^{-1}. \tag{3.6}$$

After the precoding matrix $\mathbf{D}_k$ is obtained, the perturbation vector for user $k$ can be calculated according to

$$\mathbf{l}_k = \underset{\mathbf{l}_k \in \mathbb{C}\mathbb{Z}^{n_k^r \times 1}}{\arg\min} \left\| (\mathbf{H}_k \mathbf{B}_k)^{-1} (\mathbf{s}_k + \tau_k \mathbf{l}_k) \right\|_F^2. \tag{3.7}$$

For the low-complexity BDVP method proposed in [18], the precoding matrix for user $k$ is $\mathbf{F}_k = \hat{\mathbf{H}}_k$, where $\hat{\mathbf{H}}_k$ is obtained from the decomposition of the channel's pseudo inverse $\mathbf{H}^\dagger$, given by

$$\mathbf{H}^\dagger = \left[ \hat{\mathbf{H}}_1, \hat{\mathbf{H}}_2, \cdots, \hat{\mathbf{H}}_K \right], \tag{3.8}$$

with each $\hat{\mathbf{H}}_k \in \mathbb{C}^{N_t \times n_k^r}$. The transmit signal vector for user $k$ is then expressed as

$$\mathbf{x}_k = \frac{1}{f_{VP}} \cdot \hat{\mathbf{H}}_k \left( \mathbf{s}_k + \tau_k \mathbf{l}_k \right), \tag{3.9}$$

where a unit transmit power is assumed. $\mathbf{l}_k$ is the perturbation vector for user $k$ and we obtain $\mathbf{l}_k$ following

$$\mathbf{l}_k = \underset{\mathbf{l}_k \in \mathbb{C}\mathbb{Z}^{n_k^r \times 1}}{\arg\min} \left\| \hat{\mathbf{H}}_k \left( \mathbf{s}_k + \tau_k \mathbf{l}_k \right) \right\|_F^2. \tag{3.10}$$

In (3.9), $f_{VP}$ is the transmit scaling factor that ensures $\mathbb{E}\left\{ \|\mathbf{x}\|_F^2 \right\} = 1$, given by

$$f_{VP} = \left\| \mathbf{H}^\dagger (\mathbf{s} + \tau \mathbf{l}) \right\|_F, \tag{3.11}$$

where $\mathbf{l} = \left[ \mathbf{l}_1^T, \mathbf{l}_2^T, \cdots, \mathbf{l}_K^T \right]^T$. At the receiver side, the signal vector is firstly rescaled to eliminate the effect of $f_{VP}$, and then fed to a modulo operation to remove the perturbation vector $\mathbf{l}_k$ for each user, as shown in Chapter 2.2.2.

#### 3.1.2.3 UGVP Precoding

For UGVP precoding, the difference compared to BDVP is that UGVP divides the total $K$ users into groups, such that the modulation type within each group is identical. With this approach, the sphere search for the perturbation vectors can be performed independently for each group. We denote $B$ as the total number of groups which is equal to the number of modulation types employed by the total $K$ users, and denote $\{\mathscr{G}_1, \mathscr{G}_2, ..., \mathscr{G}_B\}$ as the set of groups. For each group $\mathscr{G}_b$, we have $\tau_k^b = \tau^b$, $\forall k \in \mathscr{G}_b$, and the perturbation vector for the group $\mathscr{G}_b$ is given by

$$\mathbf{l}_b = \underset{\mathbf{l}_b \in \mathbb{C}\mathbb{Z}^{n_b \times 1}}{\arg\min} \left\| \hat{\mathbf{H}}_b \left( \mathbf{s}_b + \tau_b \mathbf{l}_b \right) \right\|_F^2, \tag{3.12}$$

where $\mathbf{H}^\dagger$ is divided by groups into $\mathbf{H}^\dagger = \left[ \hat{\mathbf{H}}_1, \hat{\mathbf{H}}_2, ..., \hat{\mathbf{H}}_B \right]$ for UGVP. $\hat{\mathbf{H}}_b \in \mathbb{C}^{N_t \times n_b}$ is the precoding matrix for the users within the group $\mathscr{G}_b$, $\mathbf{s}_b = \left[ \mathbf{s}_1^T, \mathbf{s}_2^T, ..., \mathbf{s}_{b_M}^T \right]^T$, $n_b = \sum_{k=b_1}^{b_M} n_k^r$ is the total number of transmit symbols in group $b$, and $b_M$ is the total number of users in group $b$, which leads to $N_r = \sum_{b=1}^{B} n_b$. By reducing the number of individual sphere searches required, UGVP is shown to achieve an improved performance over BDVP [18].

### 3.1.3 Proposed Joint VP Precoding

Single-modulation VP approaches search over the entire $N_r$ dimensions to obtain the perturbation vector that minimises the transmit signal norm [17]. Nevertheless, due to the different values of $\tau$ for different modulation types, conventional VP techniques cannot be directly employed in multiple-modulation scenarios. For BDVP, the sphere search is performed within $n_k^r$ dimensions for each user $k$ and a total number of $K$ independent sphere searches are conducted to obtain the perturbation vectors. Similarly, UGVP performs an $n_b$-dimensional sphere search for $B$ modulation types to obtain the desired perturbation vectors. By rewriting $\mathbf{H}^\dagger = \left[ \hat{\mathbf{H}}_1, \hat{\mathbf{H}}_2, ..., \hat{\mathbf{H}}_K \right]$ and

defining a diagonal matrix $\boldsymbol{\tau} = diag\{\tau_1 \cdot \mathbf{1}, \tau_2 \cdot \mathbf{1}, \cdots, \tau_k \cdot \mathbf{1}\}$, we obtain

$$\begin{aligned} \left\| \mathbf{H}^{\dagger} (\mathbf{s} + \boldsymbol{\tau}\mathbf{l}) \right\|_F^2 &= \left\| \sum_{k=1}^{K} \hat{\mathbf{H}}_k (\mathbf{s}_k + \tau_k \mathbf{l}_k) \right\|_F^2 \\ &\leq \sum_{k=1}^{K} \left\| \hat{\mathbf{H}}_k (\mathbf{s}_k + \tau_k \mathbf{l}_k) \right\|_F^2. \end{aligned} \tag{3.13}$$

It is found that with the increase in the value of $K$, the gap between $\left\| \mathbf{H}^{\dagger} (\mathbf{s} + \boldsymbol{\tau}\mathbf{l}) \right\|_F^2$ and $\sum_{k=1}^{K} \left\| \hat{\mathbf{H}}_k (\mathbf{s}_k + \tau_k \mathbf{l}_k) \right\|_F^2$ is larger [18]. Therefore, while UGVP performs better than BDVP, they are both sub-optimal as each sphere search is conducted in a reduced dimension.

In this section, we propose a joint perturbation technique via a constellation scaling approach, which searches the perturbation vectors in a full dimension. To be specific, without loss of generality we assume that multiple modulation types are applied, and each modulation is denoted as $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$,...,$\mathscr{N}$. We denote $N_i$ as the number of symbols that is from the constellation of the modulation type $i$, and we obtain $\sum_{i=\mathscr{A}}^{\mathscr{N}} N_i = N_r$. We first rewrite

$$\mathbf{H}^{\dagger} = [\mathbf{H}_{\mathscr{A}}, \cdots, \mathbf{H}_{\mathscr{N}}], \ \mathbf{s} + \boldsymbol{\tau}\mathbf{l} = [(\mathbf{s}_{\mathscr{A}} + \tau_{\mathscr{A}} \mathbf{l}_{\mathscr{A}})^T, \cdots, (\mathbf{s}_{\mathscr{N}} + \tau_{\mathscr{N}} \mathbf{l}_{\mathscr{N}})^T]^T, \tag{3.14}$$

where $\mathbf{s}_i$ denotes the symbol vector from the constellation of modulation $i$ and $\mathbf{H}_i$ is the corresponding channel matrix. Based on (3.14), the transmit signal norm can be reformulated as

$$\begin{aligned} \left\| \mathbf{H}^{\dagger} (\mathbf{s} + \boldsymbol{\tau}\mathbf{l}) \right\|_F^2 &= \| [\mathbf{H}_{\mathscr{A}}, \cdots, \mathbf{H}_{\mathscr{N}}] \cdot (\mathbf{s} + \boldsymbol{\tau}\mathbf{l}) \|_F^2 \\ &= \left\| \sum_{i=\mathscr{A}}^{\mathscr{N}} \frac{\tau_i}{\tau_0} \cdot \frac{\tau_0}{\tau_i} \mathbf{H}_i (\mathbf{s}_i + \tau_i \mathbf{l}_i) \right\|_F^2 \\ &= \left\| \sum_{i=\mathscr{A}}^{\mathscr{N}} \frac{\tau_i}{\tau_0} \mathbf{H}_i (\frac{\tau_0}{\tau_i} \mathbf{s}_i + \tau_0 \mathbf{l}_i) \right\|_F^2, \end{aligned} \tag{3.15}$$

where $\mathbf{l}_i$ is the perturbation vector for modulation type $i$. $\tau_0$ denotes the modulo basis that is employed in the sphere search. For our proposed approach,

it is easy to observe that $\tau_0$ can be an arbitrary positive value. By defining

$$\mathbf{G}=\left[\frac{\tau_{\mathscr{A}}}{\tau_0}\mathbf{H}_{\mathscr{A}},\cdots,\frac{\tau_{\mathscr{N}}}{\tau_0}\mathbf{H}_{\mathscr{N}}\right],\ \mathbf{t}=\left[\frac{\tau_0}{\tau_{\mathscr{A}}}s^T_{\mathscr{A}},\cdots,\frac{\tau_0}{\tau_{\mathscr{N}}}s^T_{\mathscr{N}}\right]^T, \tag{3.16}$$

(3.15) can be further transformed into

$$\left\|\mathbf{H}^{\dagger}\left(\mathbf{s}+\boldsymbol{\tau}\mathbf{l}\right)\right\|_F^2=\left\|\mathbf{G}\left(\mathbf{t}+\tau_0\mathbf{l}\right)\right\|_F^2, \tag{3.17}$$

which is in a standard form of the conventional single-modulation VP precoding as in [17].

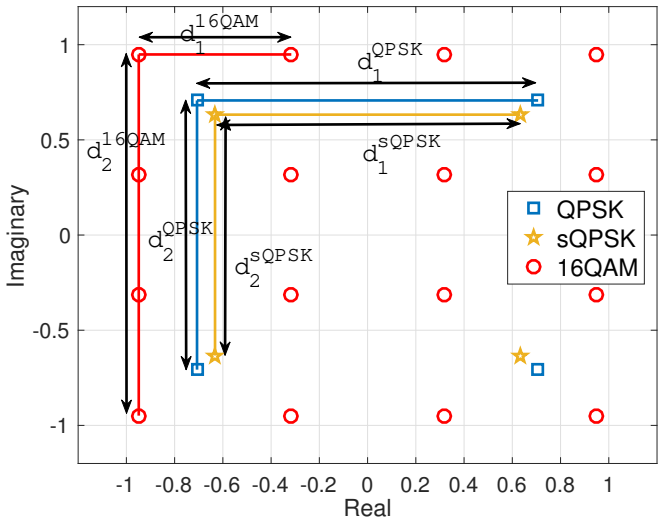



**Figure 3.1:** Normalised constellation points for QPSK, scaled QPSK and 16QAM

With the above transformation, different modulations share an identical modulo basis $\tau_0$, and the sphere search can be conducted jointly as for conventional VP. Accordingly, no performance losses will be observed compared to the conventional VP for single modulation. The proposed method can also be interpreted as follows: by scaling each constellation of different modulations and the corresponding channel matrices, the modulo basis $\tau$ for each modulation is made identical, and therefore the search for the perturbation vector can be performed jointly. To illustrate how the proposed scheme works, Fig. 3.1 depicts the normalised constellation points of 16QAM, QPSK,

and QPSK after the proposed constellation scaling approach (denoted as sQPSK). According to the definition of $\tau$ [17], in Fig. 3.1 we can obtain the values of the modulo basis $\tau_{16QAM} = 2.5298$ for 16QAM and $\tau_{QPSK} = 2.8284$ for QPSK respectively, shown before (2.23) in Chapter 2.2.2. We scale the constellation of QPSK by $\tau_{16QAM}/\tau_{QPSK}$ such that

$$\frac{\tau_{16QAM}}{\tau_{sQPSK}} = \frac{d_1^{16QAM} + d_2^{16QAM}}{d_1^{sQPSK} + d_2^{sQPSK}} = 1, \tag{3.18}$$

i.e., the value of the modulo basis for the scaled QPSK is equal to that of 16QAM. With the above constellation scaling, $\tau_0 = \tau_{16QAM}$ and the joint perturbation can be performed.

### 3.1.4 Computational Cost Analysis

In this section, we compare the computational cost of the proposed method with BDVP and UGVP in terms of the floating-point operations required [87, 88]. The inversion of the $n_k \times n_k$ matrix $\mathbf{D}_k$ and $N_r \times N_t$ matrix $\mathbf{H}$ using Gauss-Jordan elimination requires $4n_k^3/3$ and $4N_t^3/3$ operations, respectively. An $n$-dimension sphere search requires $\mathcal{O}\{n^6\}$ operations [88], where $\mathcal{O}\{\cdot\}$ defines the order of numerical operators. In addition, the calculation of $\mathbf{G}$ and $\mathbf{t}$ involved in the JVP precoding requires $\mathcal{O}\{N_t\}$ operations. This additional computational cost contributes little to the total complexity, compared to that of the sphere search. Accordingly, the proposed algorithm almost consumes the same complexity as the conventional VP approach, which will be further shown numerically. Subsequently, the computational cost for each VP-based methods can be calculated as

$$\begin{aligned}
\mathrm{C_{BDVP}} &= \mathcal{O}\{4N_t^3/3\} + \sum_{k=1}^{K} \mathcal{O}\{n_k^6\} = \mathcal{O}\left\{4N_t^3/3 + \sum_{k=1}^{K} n_k^6\right\}, \\
\mathrm{C_{UGVP}} &= \mathcal{O}\{4N_t^3/3\} + \sum_{b=1}^{B} \mathcal{O}\{n_b^6\} = \mathcal{O}\left\{4N_t^3/3 + \sum_{b=1}^{B} n_b^6\right\}, \\
\mathrm{C_{JVP}} &= \mathcal{O}\{4N_t^3/3\} + \mathcal{O}\{N_t^6\} + \mathcal{O}\{2N_t\} = \mathcal{O}\left\{4N_t^3/3 + N_t^6 + 2N_t\right\}.
\end{aligned} \tag{3.19}$$

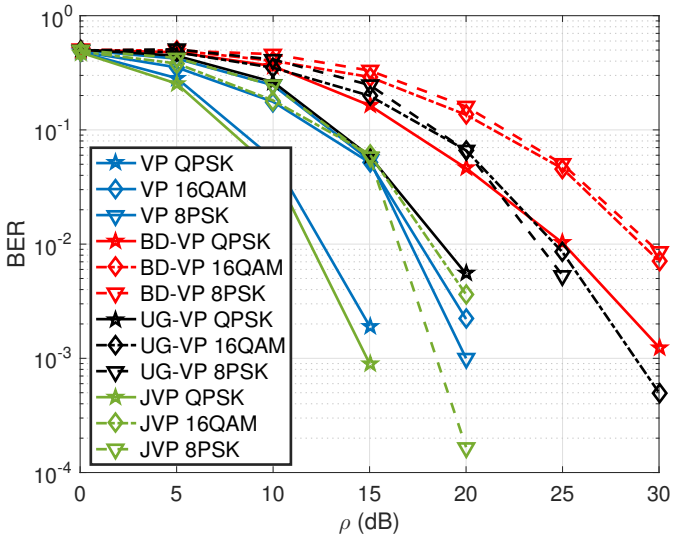



**Figure 3.2:** BER v.s. transmit SNR $\rho$, $N_t = 12$, $K$=6, $n_k^r = 2$, $\forall k$

### 3.1.5 Numerical Results

In this section, the performance of the proposed JVP is compared to existing the BDVP and UGVP methods by Monte Carlo simulations. The transmit SNR is defined as $\rho = 1/\sigma^2$, and we assume the number of transmit antenna $N_t = 12$, the total number of users $K = 6$, and $n_k^r = 2$ for each user. In the existence of multiple modulations, QPSK, 8PSK, and 16QAM modulations are employed as examples, with 2 users for each modulation. The above parameters remain constant throughout the simulations, unless otherwise stated. We note that our interest lies in the multiple-modulation scenarios, and accordingly our performance comparison is the proposed 'JVP' approach with the existing 'BDVP' and 'UGVP', while the performance of the conventional VP for single modulation is only presented as reference.

In Fig. 3.2, the BER performance of the proposed JVP is compared with existing BDVP, UGVP and conventional single-modulation VP with respect to the increasing transmit SNR. It can be observed that the proposed JVP achieves a significant performance gain over both the BDVP and UGVP precoding, where a comparable performance to the conventional VP approach for each modulation is achieved. It is worth noting that the performance of PSK modulations in JVP precoding achieves an even better performance

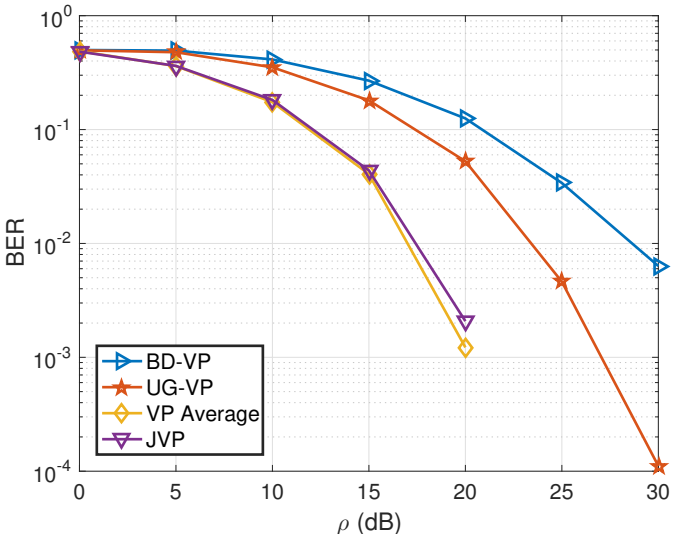



**Figure 3.3:** Average BER v.s. transmit SNR $\rho$, $N_t$=12, $K$=6, $n_k^r = 2, \forall k$

than in the conventional VP approach, as the proposed JVP approach obtains a smaller scaling factor $f_{VP}$ than the single-modulation VP approach for QPSK or 8PSK.

To further reveal the fact that the proposed JVP precoding achieves a similar performance to the conventional VP method for single modulation, we compare the average BER performance of BDVP, UGVP, JVP with the conventional VP over all modulation types in Fig. 3.3, where 'VP Average' denotes the average BER performance of the conventional VP approach with QPSK, 8PSK, and 16QAM, which is defined as

$$BER_{average} = \frac{N_{errors}^{QPSK} + N_{errors}^{8PSK} + N_{errors}^{16QAM}}{N_{bits}^{QPSK} + N_{bits}^{8PSK} + N_{bits}^{16QAM}}, \tag{3.20}$$

where $N_{bits}^{\mathscr{U}}$ denotes the total number of bits transmitted for the modulation type $\mathscr{U}$, while $N_{errors}^{\mathscr{U}}$ denotes the total number of erroneous bits in $N_{bits}^{\mathscr{U}}$. In Fig. 3.3, a similar trend to Fig. 3.2 can be observed, and it can be seen that UGVP outperforms BDVP, while the proposed JVP and 'VP Average' achieve a similar BER performance and both outperform BDVP and UGVP.

Fig. 3.4 compares the computational costs of different VP methods with respect to the number of transmit antennas, where $N_r = N_t$. It is not surpris-

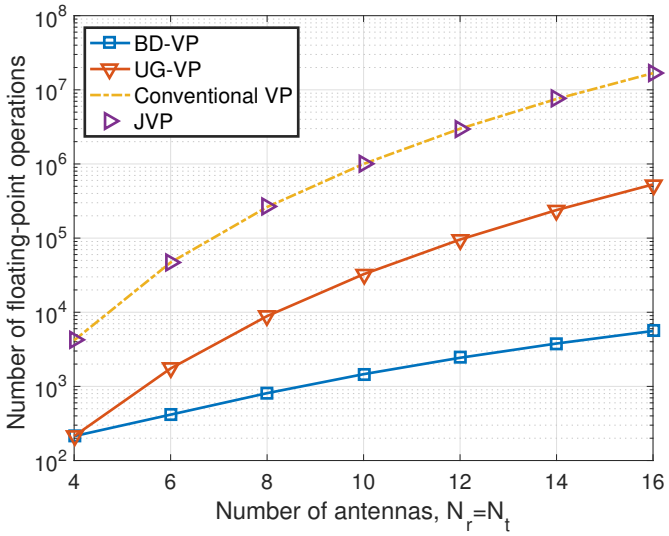



**Figure 3.4:** Computational cost v.s. number of antennas, $N_t = N_r$

ing to observe that the computational costs of all VP-based approaches increase with the increasing number of data streams. Due to a reduced search dimension compared to conventional VP, the computational cost of BDVP and UGVP is lower, while as expected the complexity of the proposed JVP is comparable to that of the conventional VP, because the computational cost is dominated by the sphere search dimension, shown in (3.19).

## 3.2 Joint Constructive Vector Perturbation

### 3.2.1 Introduction

While VP-based precoding techniques can achieve a significant performance gain over linear ZF-based approaches, the computational cost of VP is usually very high due to the sophisticated sphere search algorithms. In order to make VP techniques realisable in practice, alternative VP techniques have been investigated in order to alleviate the high computational burden [36, 37, 38]. In [36], a thresholded VP technique is proposed where the search for the perturbation vectors is terminated when a pre-defined threshold is reached. A selective perturbation is further proposed in [37] by only perturbing a part of the transmit symbols, where a flexible tradeoff between performance and

complexity is observed. In [38], the perturbation operation is replaced by a symbol-scaling approach based on the CI formulation introduced in [54, 55], where the sphere search is replaced by symbol scaling that can be efficiently obtained by solving a quadratic programming (QP) optimisation, which significantly reduces the computational complexity.

In this section, we first introduce a JCVP approach for PSK modulations and show that the performance of VP can be further improved, without incurring significant additional computational cost. By introducing a symbol-scaling vector upon the conventional VP approach and jointly optimising both the perturbation vector and the symbol-scaling vector, the noise amplification effect is minimised and at the same time the transmit symbols are enhanced by the symbol-scaling vector, which therefore leads to an improved detection performance. We further extend the proposed JCVP to multiple-modulation scenarios by replacing the perturbation for PSK users with the symbol-scaling approach while retaining the conventional VP operation for QAM users. It will be shown that the proposed method can achieve a comparable performance to JVP, while providing a significant gain in the computational cost over the JVP proposed in the previous section.

### 3.2.2 System Model

In this section, a similar MU-MIMO system downlink is considered as in Chapter 3.1, where the BS with $N_t$ transmit antennas communicates with $K$ multi-antenna users simultaneously. The number of receive antennas for each user $k$ is $n_k^r$. The received signal vector can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{3.21}$$

where $\mathbf{x}$ is the precoded signal vector to be transmitted.

### 3.2.3 Proposed JCVP for PSK

We first expand the channel's pseudo inverse matrix and the transmit symbols into their real part and imaginary part, expressed as

$$\bar{\mathbf{H}}^{\dagger} = \begin{bmatrix} \Re\left(\mathbf{H}^{\dagger}\right) & -\Im\left(\mathbf{H}^{\dagger}\right) \\ \Im\left(\mathbf{H}^{\dagger}\right) & \Re\left(\mathbf{H}^{\dagger}\right) \end{bmatrix}, \ \bar{\mathbf{s}} = \left[\Re(\mathbf{s})^{T}, \Im(\mathbf{s})^{T}\right]^{T}. \tag{3.22}$$

To exploit the CI, we follow the symbol-scaling formulation in Chapter 2.2.4 and formulate a diagonal symbol-scaling matrix $\bar{\mathbf{\Phi}} \in \mathbb{R}^{2N_r \times 2N_r}$ is applied upon the VP approach, and we construct the transmit signal vector as

$$\bar{\mathbf{x}} = \frac{1}{f_{VP}^{J}} \cdot \bar{\mathbf{H}}^{\dagger} \bar{\mathbf{\Phi}} \left(\bar{\mathbf{s}} + \tau \cdot \bar{\mathbf{l}}\right), \tag{3.23}$$

where $\bar{\mathbf{l}} \in \mathbb{Z}^{2N_r \times 1}$ is the perturbation vector obtained from standard sphere search algorithms, which is expanded by the real and imaginary part. The scaling factor $f_{VP}^{J}$ represents the noise amplification effect for the proposed precoder, which is obtained as

$$f_{VP}^{J} = \left\| \bar{\mathbf{H}}^{\dagger} \bar{\mathbf{\Phi}} \left(\bar{\mathbf{s}} + \tau \cdot \bar{\mathbf{l}}\right) \right\|_{F}. \tag{3.24}$$

For the proposed precoder, the optimal $\bar{\mathbf{\Phi}}$ is therefore the one that minimises the scaling factor $f_{VP}^{J}$, and we construct the optimisation problem on $\bar{\mathbf{\Phi}}$ as

$$\begin{aligned} \mathscr{P}_{3.1}: \ & \min_{\bar{\mathbf{\Phi}}} \left\| \bar{\mathbf{H}}^{\dagger} \bar{\mathbf{\Phi}} \left(\bar{\mathbf{s}} + \tau \cdot \bar{\mathbf{l}}^{*}\right) \right\|_{F}^{2} \\ & s.t. \quad \varphi_n \geq 1, \ \forall n \in \{1, 2, ..., 2N_t\} \end{aligned} \tag{3.25}$$

where $\varphi_n = \left[\bar{\mathbf{\Phi}}\right]_{n,n}$ is the $n$-th diagonal element in $\bar{\mathbf{\Phi}}$. By denoting $\bar{\mathbf{D}} = diag\left(\bar{\mathbf{s}} + \tau \cdot \bar{\mathbf{l}}^{*}\right)$ and rearranging the matrix expression with the vectorisation

operation, the minimisation problem can be further transformed into

$$\begin{aligned} \mathscr{P}_{3.2}: \quad & \min_{\bar{\mathbf{\Phi}}} \left\| \bar{\mathbf{H}}^{\dagger} \bar{\mathbf{D}} \cdot \mathrm{vec}\left(\bar{\mathbf{\Phi}}\right) \right\|_F^2 \\ s.t. \quad & \varphi_n \geq 1, \, \forall n \in \{1, 2, ..., 2N_t\} \end{aligned} \tag{3.26}$$

$\mathscr{P}_{3.2}$ is a standard QP optimisation that can be efficiently solved with interior-point methods. Finally, the signal vector to be transmitted after precoding can be expressed as

$$\mathbf{x} = \frac{1}{f_{VP}^J} \cdot \mathbf{H}^{\dagger} \mathbf{\Phi} \left(\mathbf{s} + \tau \cdot \mathbf{l}\right), \tag{3.27}$$

where $\mathbf{\Phi}\left(\mathbf{s} + \tau \cdot \mathbf{l}\right)$ is the complex equivalence of $\bar{\mathbf{\Phi}}\left(\bar{\mathbf{s}} + \tau \cdot \bar{\mathbf{l}}\right)$. At the receiver, the received signal vector can be obtained as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \frac{1}{f_{VP}^J} \cdot \left(\mathbf{\Phi}\mathbf{s} + \tau \cdot \mathbf{\Phi}\mathbf{l}\right) + \mathbf{n}. \tag{3.28}$$

It is observed from (3.28) that the perturbation vector for the proposed precoder becomes $\tau \cdot \mathbf{\Phi}\mathbf{l}$ instead of $\tau \cdot \mathbf{l}$ for conventional VP methods due to the scaling of the transmit symbols. Therefore, for each user $k$, the modulo basis that is used to remove the perturbation should also be scaled by $\mathbf{\Phi}$, and the output of the modulo operation for the $m$-th stream can be obtained as

$$\begin{aligned} r_m &= \mathrm{mod}\left[f_{VP}^J \cdot y_m\right] \\ &= \mathrm{mod}\left[\Phi_m s_m + \Phi_m \tau \cdot l_m + f_{VP}^J \cdot n_m\right] \\ &= s_m + \tilde{n}_m, \end{aligned} \tag{3.29}$$

where $\Phi_m = [\mathbf{\Phi}]_{m,m}$ is the $m$-th diagonal element in matrix $\mathbf{\Phi}$. The modulo function $\mathrm{mod}\left[x_m\right]$ for the proposed JCVP is reconstructed as

$$\mathrm{mod}\left[x_m\right] = f_{\tau \cdot \varphi_m}\left\{\Re\left(x_m\right)\right\} + j \cdot f_{\tau \cdot \varphi_{\hat{m}}}\left\{\Im\left(x_m\right)\right\}, \tag{3.30}$$

where $\hat{m} = m + \frac{N_t}{2}$. $f_{\tau \cdot \varphi_m}(x)$ in (3.30) is given by

$$f_{\tau \cdot \varphi_m}(x) = x - \left\lfloor \frac{x + \varphi_m \tau / 2}{\varphi_m \tau} \right\rfloor \cdot \varphi_m \tau. \tag{3.31}$$

**Remark**: While the proposed JCVP will be shown to achieve an improved performance over the conventional VP technique for PSK modulations, the proposed JCVP is however difficult to implement in practical wireless systems, because the modulo function employed at the receiver is related to the symbol-scaling vector that is dependent on the transmit symbols, shown in (3.31). Nevertheless, the JCVP method introduced above motivates a more practical JCVP method for multiple-modulation cases, where the modulo function for PSK modulations is no longer necessary. This circumvents the above drawbacks, which is introduced in the following.

### 3.2.4 Proposed JCVP for Multiple-Modulation Scenarios

The proposed JCVP does not directly apply to multiple-modulation scenarios, because the symbol-scaling approach for QAM users will lead to incorrect demodulation. To overcome this, in this section we propose to employ a partial symbol-scaling approach for PSK users and a partial VP approach for QAM users [37]. To be more specific, for QAM users, the conventional VP approach is employed and the perturbation vector can be obtained as

$$\mathbf{l} = \arg\min \left\| \hat{\mathbf{H}}^{\dagger} \left( \hat{\mathbf{s}} + \tau \cdot \begin{bmatrix} \mathbf{0}_{N_t - L} \\ \mathbf{l}_L \end{bmatrix} \right) \right\|_F^2 = \arg\min \left\| \hat{\mathbf{H}}^{\dagger} \hat{\mathbf{s}} + \tau \cdot \hat{\mathbf{H}}_L^{\dagger} \mathbf{l}_L \right\|_F^2, \tag{3.32}$$

where $L$ is the number of QAM users. $\hat{\mathbf{H}}^{\dagger} = \left[ \hat{\mathbf{H}}_{N_t - L}^{\dagger}, \hat{\mathbf{H}}_L^{\dagger} \right]$ and $\hat{\mathbf{s}}$ denote the appropriate reordering of $\mathbf{H}^{\dagger}$ and $\mathbf{s}$. With this transformation, only the transmit symbols for users that employ QAM modulations are perturbed. By further denoting

$$\mathbf{u} = \hat{\mathbf{s}} + \tau \cdot \begin{bmatrix} \mathbf{0}_{N_t - L} \\ \mathbf{l}_L^* \end{bmatrix} = \left[ \hat{\mathbf{s}}_{N_t - L}^T, (\hat{\mathbf{s}}_L + \tau \cdot \mathbf{l}_L^*)^T \right]^T, \tag{3.33}$$

we employ a symbol-scaling approach for PSK symbols only, and the precoded signals are given by

$$\hat{\bar{\mathbf{x}}} = \frac{1}{\hat{f}_{VP}^{J}} \cdot \hat{\bar{\mathbf{H}}}^{\dagger} \hat{\bar{\mathbf{\Phi}}} \bar{\mathbf{u}}, \tag{3.34}$$

where $\hat{f}_{VP}^{J} = \left\| \hat{\bar{\mathbf{H}}}^{\dagger} \hat{\bar{\mathbf{\Phi}}} \bar{\mathbf{u}} \right\|_F^2$, $\hat{\bar{\mathbf{H}}}^{\dagger}$ and $\bar{\mathbf{u}}$ denote the expanded $\hat{\mathbf{H}}^{\dagger}$ and $\mathbf{u}$ by the real and imaginary part, respectively. Subsequently, by applying a similar approach as in Chapter 3.2.3, the optimal symbol-scaling factor $\hat{\bar{\mathbf{\Phi}}}$ for multiple-modulation cases can be obtained by solving the following QP optimisation

$$\begin{aligned} \mathscr{P}_{3.3}: \quad & \min_{\hat{\bar{\mathbf{\Phi}}}} \left\| \hat{\bar{\mathbf{H}}}^{\dagger} \hat{\bar{\mathbf{D}}} \cdot \mathrm{vec}\left( \hat{\bar{\mathbf{\Phi}}} \right) \right\|_F^2 \\ s.t. \quad & \hat{\varphi}_n \geq 1, \ \forall n \in \{1, 2, ..., N_r - L, N_r + 1, N_r + 2, ..., 2N_r - L\} \\ & \hat{\varphi}_n = 1, \ \forall n \in \{N_r - L + 1, ..., N_r, 2N_r - L + 1, ..., 2N_r\} \end{aligned} \tag{3.35}$$

where $\hat{\varphi}_k$ denotes the $k$-th element in $\hat{\bar{\mathbf{\Phi}}}$ and $\hat{\bar{\mathbf{D}}} = diag(\bar{\mathbf{u}})$. The second constraint ensures that the symbols from QAM modulations are not scaled in order to avoid detection errors. Finally, the precoded and perturbed signal vector to be transmitted can be expressed as

$$\hat{\mathbf{x}} = \frac{1}{\hat{f}_{VP}^{J}} \cdot \hat{\mathbf{H}}^{\dagger} \hat{\mathbf{\Phi}} \left( \hat{\mathbf{s}} + \tau \begin{bmatrix} \mathbf{0}_{N_t - L} \\ \mathbf{l}_L \end{bmatrix} \right), \tag{3.36}$$

and the scaling factor $\hat{f}_{VP}^{J}$ in the transmit signal vector is accordingly obtained as

$$\begin{aligned} \hat{f}_{VP}^{J} &= \left\| \hat{\mathbf{H}}^{\dagger} \hat{\mathbf{\Phi}} \mathbf{u} \right\|_F \\ &= \left\| \left[ \hat{\mathbf{H}}_{N_t - L}^{\dagger}, \hat{\mathbf{H}}_{L}^{\dagger} \right] \cdot \begin{bmatrix} \mathbf{\Phi}_{N_t - L} \cdot \hat{\mathbf{s}}_{N_t - L} \\ \hat{\mathbf{s}}_L + \tau \cdot \mathbf{l}_L \end{bmatrix} \right\|_F \\ &= \left\| \hat{\mathbf{H}}_{N_t - L}^{\dagger} \mathbf{\Phi}_{N_t - L} \hat{\mathbf{s}}_{N_t - L} + \hat{\mathbf{H}}_{L}^{\dagger} (\hat{\mathbf{s}}_L + \tau \cdot \mathbf{l}_L) \right\|_F. \end{aligned} \tag{3.37}$$

**Remark 1**: It can be seen that the sequence of the sphere search process and the symbol-scaling process can also be employed conversely. In the following section, it will be numerically shown that performing the sphere

search first will always lead to an improved performance, which is due to the fact that in the second process, compared to conventional VP that searches only limited integers, the symbol-scaling approach searches the entire constructive areas such that there is a higher possibility that the noise amplification effect is minimised.

**Remark 2**: While only two modulation types are considered in the above derivation, the extension to multiple modulations is trivial. By applying JVP introduced in Chapter 3.1 first, multiple modulation types can be divided into PSK-based modulations and QAM-based modulations. Accordingly, the symbol-scaling approach can be applied to PSK-based symbols while the perturbation operation can be applied to QAM-based symbols.

### 3.2.5 Computational Cost Analysis

Following the complexity analysis in Chapter 3.1.4, the complexity of the proposed schemes for PSK modulations can be expressed as

$$C_{JCVP} = \mathscr{O}\left\{4N_t^3/3\right\} + \mathscr{O}\left\{N_t^6\right\} + \mathscr{O}\left\{8N_t^3T\right\} = \mathscr{O}\left\{4N_t^3/3 + N_t^6 + 8N_t^3T\right\}, \tag{3.38}$$

where $\mathscr{O}\left\{8N_t^3T\right\}$ is the complexity of a QP optimisation with $M$ dimensions, and $T$ is given by [38]

$$T = \log_2\left(\max\left|\delta_{i,j}\right| + 1\right) + \log_2\left(4M\right) + 1, \tag{3.39}$$

where $\delta_{i,j}$ is the $i,j$-th element of the matrix $\mathbf{\Omega}^{1/2}$. For PSK modulations, $\mathbf{\Omega}^{1/2} = \bar{\mathbf{H}}^{\dagger}\bar{\mathbf{D}}$ and $\mathbf{\Omega}^{1/2} = \hat{\bar{\mathbf{H}}}^{\dagger}\hat{\bar{\mathbf{D}}}$ for the cases of multiple modulations. Similarly, the complexity of the proposed scheme can be obtained as

$$\begin{aligned} C_{JCVP-AM} &= \mathscr{O}\left\{4N_t^3/3\right\} + \mathscr{O}\left\{L^6\right\} + \mathscr{O}\left\{8(N_t-L)^3T\right\} \\ &= \mathscr{O}\left\{4N_t^3/3 + L^6 + 8(N_t-L)^3T\right\}. \end{aligned} \tag{3.40}$$

### 3.2.6 Numerical Results

In this section, Monte Carlo simulations are utilised to evaluate the performance of the proposed schemes. We assume $N_t = 8$, and a total number of $K = 4$ users with $n_k^r = 2$ for each user. We consider both QPSK scenarios and multiple-modulation scenarios where 2 users apply QPSK and 2 users apply 16QAM. For the case of multiple-modulation scenarios, we denote 'CVP-VP' as the scheme where symbol-scaling approach is performed first and 'VP-CVP' as the scheme where sphere search is performed first.

Fig. 3.5 compares the BER performance of the proposed JCVP with the conventional VP in [17], BDVP in [48] and CVP in [38] with respect to the transmit SNR for QPSK. As can be observed, the proposed JCVP further improves the performance of conventional VP and achieves the best performance among all techniques. It is also observed that CVP can achieve a better performance when the transmit SNR is low, as it exploits the CI to benefit its detection performance, while the performance advantage diminishes with the increase in the transmit SNR.

In Fig. 3.6, the BER performance of the proposed schemes is shown in multiple-modulation scenarios with respect to the transmit SNR. For users

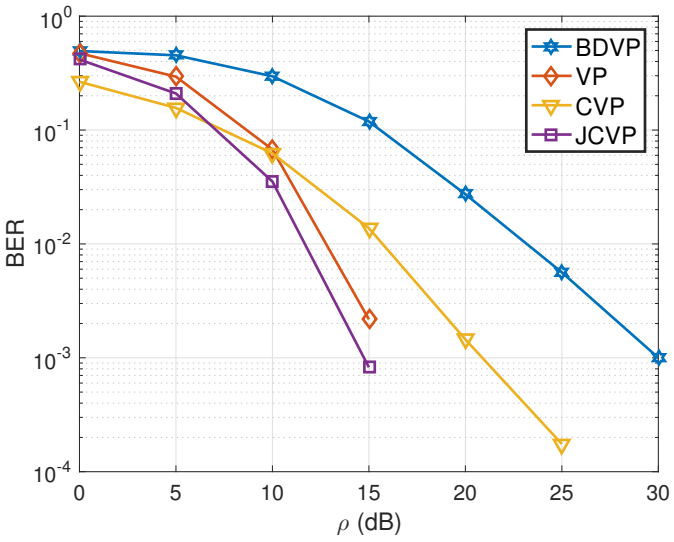



**Figure 3.5:** BER v.s. transmit SNR $\rho$, $N_t = 8$, $K$=4, $n_k^r = 2$, $\forall k$, QPSK.

employing QPSK, it is shown that the proposed VP-CVP improves the performance of JVP and achieves the best performance, as the transmit symbols benefit from the symbol-scaling approach. For users employing 16QAM, the proposed VP-CVP performs slightly worse than JVP as the sphere search is performed within a smaller dimension. It is also observed that VP-CVP outperforms CVP-VP, for both QPSK and 16QAM.

In Fig. 3.7, the computational cost is shown for both QPSK and multiple-modulation scenarios where both QPSK and 16QAM are employed, where 'JCVP-MM' denotes the computational cost of the proposed JCVP technique in multiple-modulation cases. The number of users that employs either modulation type in the cases of multiple modulations is $N = K/2$. For the proposed JCVP, its computational cost is only slightly higher than that of the conventional VP, since the complexity is dominated by the sphere search algorithm. The proposed JCVP for multiple-modulation cases is much more computationally efficient than JCVP because the sphere search dimension is reduced from $K$ to $K/2$, which reveals an improved performance-complexity tradeoff for the proposed JCVP for multiple-modulation scenarios, compared to the proposed JVP in Chapter 3.1.

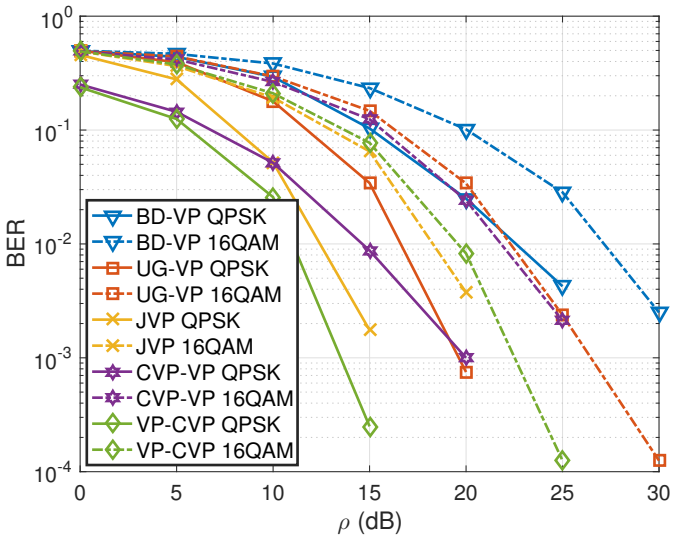



**Figure 3.6:** BER v.s. transmit SNR $\rho$, $N_t = 8$, $K = 4$, $n_k^r = 2$, $\forall k$, multiple modulations

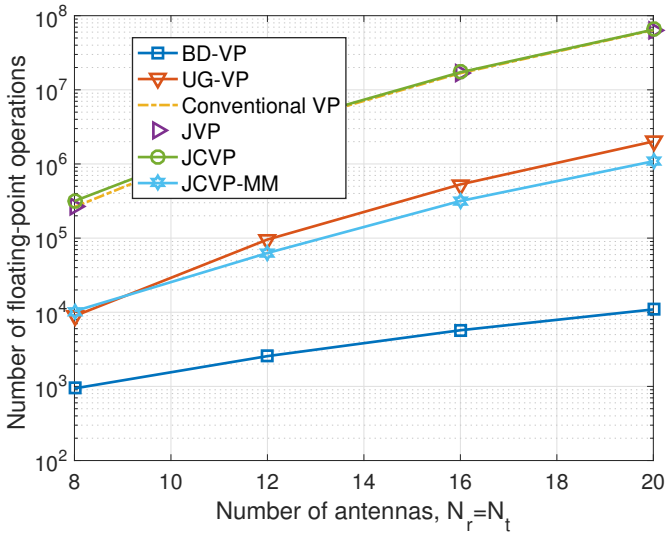



**Figure 3.7:** Computational cost v.s. number of antennas, $N_t = N_r$

## 3.3 Summary

In this chapter, we investigate the non-linear VP-based precoding techniques and propose novel approaches to further enhance the performance of existing VP approaches. The proposed JVP via the constellation scaling method is shown to outperform existing VP-based schemes and achieve a comparable performance to the conventional single-modulation VP in multiple-modulation scenarios. A JCVP approach based on CI and partial perturbation is further proposed, which achieves an improved performance-complexity tradeoff compared to the proposed JVP in multiple-modulation cases.

# Chapter 4

# Mutual Coupling Exploitation for P2P MIMO Systems

This chapter is based on our works published in [C7], [J5].

## 4.1 Introduction

Inspired by the compact parasitic-antenna arrays where the MC effect is exploited, in this chapter we discuss the exploitation of the coupling effect for generic FD compact single-user multiple-antenna systems.

Most of the existing MIMO techniques are studied under the ideal assumption that the antenna spacing between adjacent antennas in the antenna array is larger than half of the wavelength, such that the spatial correlation and coupling effect can be negligible. Nevertheless, compact multiple-antenna systems such as light-weight and small access points (APs) usually have strict size constraints. In such cases, the correlation effect and MC effect in the antenna array cannot be neglected [57, 61, 62, 91]. The effect of spatial correlation and antenna coupling is studied in [61, 62, 91], when an increasing number of antennas is fitted in a fixed physical space. The transmit correlation effect has been experimentally studied in [57, 92, 93], and how the correlation affects the performance of multiple-antenna systems has been investigated in [94, 95, 96, 97]. Further studies in [98, 99, 100] have focused on the robust precoding techniques for correlated channels.

As for the MC, which is more relevant to this chapter, in [64] its effect on a $2 \times 2$ MIMO system has been analytically studied based on the derived MC matrix. Further studies on the MC effect can be found in [63, 71, 72, 73], where the coupling effect in Rician channels are considered in [71]. In [73], it is shown that the MC has a negative effect on the rate performance in the uplink, when the antenna spacing becomes smaller. Further studies of the coupling effect on the channel capacity with respect to the antenna spacing can be found in [101, 102, 103, 104]. The above studies show that for the performance of MIMO systems, the existence of MC can be advantageous in some cases while disadvantageous in other cases.

In order to compensate for the performance loss when a negative coupling effect is observed, a number of approaches have been proposed [74, 75, 76], most of which are based on the derivation of a compensation matrix. In [74], the effectiveness of the compensation matrix is validated by experiments. In [75], the compensation matrix is derived based on a generalised scattering matrix of the antenna array and the spherical mode expansion. In [76], the MC compensation techniques are studied for both transmitting and receiving antenna arrays. Moreover, novel structures such as matching circuits are further introduced in [7, 20, 77] to manage the detrimental coupling effect. In [7], by adding parasitic elements to the antenna array, a reverse coupling effect is formulated to offset the MC effect. In [20], the impact of the coupling effect at low terahertz (THz) frequencies is studied, where a mantle cloaking method is applied for strip dipole antennas to reduce the MC effect. A novel structure is proposed in [77] to suppress the coupling effect by adding a U-shaped microstrip. Other techniques targeting at the MC compensation can be found in [19, 105, 106]. Nevertheless, it should be noted that, most of these existing approaches are not from a signal processing perspective and therefore only apply to specific antenna arrays.

Different from the above methods that are aimed at minimising or compensating for the coupling effect, in this chapter we propose to exploit the

coupling effect by the use of tunable antenna loads, which further improves the performance of the generic compact multiple-antenna systems. With the deployment of the tunable load impedance, for example a varactor, the MC effect can be controlled by changing the values of each antenna load. We first consider the full elimination of the coupling effect, where it is analytically shown that it is not achievable to fully remove the MC effect by solely tuning the values of the antenna loads. While the full elimination of the coupling effect is not possible, we proceed to study the exploitation of this effect and show that an improved performance can be obtained. By controlling each tunable load value based on the CI formulation [41], the inter-symbol interference generated by the MC effect is rotated and constructively aligned to the phase of the desired transmit symbol. It is the concept of constructive region in the CI formulation that further relaxes the feasible region for each tunable antenna load, which leads to a realisable constructive coupling matrix. The exploitation of the coupling effect is subsequently constructed as optimisation problems on the tunable antenna loads, and the practical constraints on each antenna load with respect to the antenna radiation are taken into consideration. Moreover, the analytical probability of error is studied by characterising the separately-correlated channels, and we also discuss the implementation of the proposed methods in practice, where a look-up table can be built based on the transmit symbol combination prior to data transmission, such that the proposed techniques can be efficiently applied afterwards. Our proposed approaches reveal the possibility to exploit the MC effect to further benefit the system performance instead of compensating for this effect, which is further validated by numerical results.

The contributions of this chapter are summarised as

- We propose to exploit the MC effect for P2P MIMO systems by means of analog processing with tunable antenna loads, where both PSK- and QAM-modulated data symbols are considered.
- The optimisation problems are constructed by exploiting the CI for-

mulation based on two criteria, and further transformed into convex optimisation problems, where practical constraints are considered.

- We characterise the separately-correlated channel model, and mathematically analyse the system performance in terms of the analytical probability of error.
- We discuss the practical implementation of the proposed methods, where a look-up table can be built dependent on the transmit symbol combination.

## 4.2 System Model

We consider a P2P MIMO system with $N_t$ antennas at the transmitter and $N_r$ antennas at the receiver, as shown in Fig. 4.1, where $N_r \geq N_t$. We assume a total number of $N_t$ streams, and each data symbol is formed from a normalised constellation of a PSK or QAM modulation, which is denoted as $\mathbf{s} \in \mathbb{C}^{N_t \times 1}$. For PSK modulations, each data symbol satisfies $|s_k| = 1$, where a fixed power per symbol is maintained. With the MC effect considered at the transmitter, the effective channel is composed by a coupling matrix $\mathbf{Z} \in \mathbb{C}^{N_t \times N_t}$, followed by the propagation channel $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$. The received signal vector can be obtained as

$$\mathbf{y} = \mathbf{HZs} + \mathbf{n}, \tag{4.1}$$

where $\mathbf{n}$ is the additive Gaussian noise vector with $\mathbf{n} \sim \mathbb{CN}\left(\mathbf{0}, \sigma^2 \cdot \mathbf{I}\right)$. Throughout this chapter, we assume full CSI at the receiver side, and the

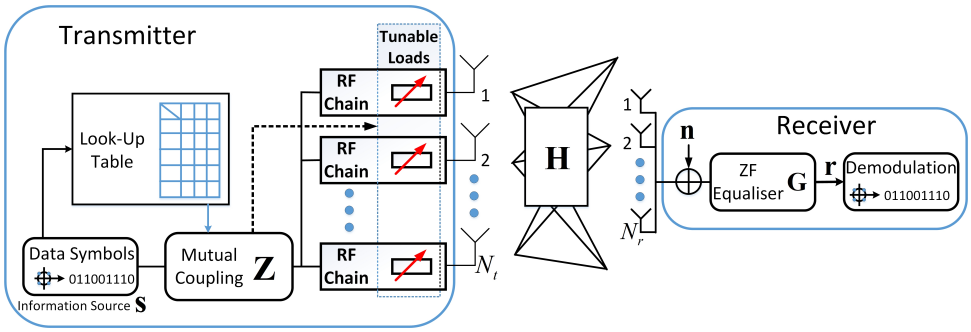



**Figure 4.1:** System model of a P2P MIMO system with MC exploitation

discussion on the CSI acquisition is briefly presented in Chapter 4.5.1. We focus on the MC exploitation at the transmitter with tunable antenna loads, while we assume the antenna spacing for the antenna array at the receiver is large, i.e., there is no MC effect at the receiver.

**Remark:** In the case where the receiver also employs a compact array, there exists coupling effect for the receive antenna array. Then, an effective channel $\hat{\mathbf{H}} = \mathbf{Z}_r\mathbf{H}$ can be constructed to incorporate the fixed coupling effect $\mathbf{Z}_r$ at the receiver into the channel. This does not affect the exploitation of the coupling effect at the transmitter.

At the receiver, the signal vector is linearly filtered by the equaliser $\mathbf{G} \in \mathbb{C}^{N_t \times N_r}$, and the estimation for the intended data vector can be expressed as

$$\mathbf{r} = \mathbf{G}\mathbf{y} = \mathbf{G}\mathbf{H}\mathbf{Z}\mathbf{s} + \mathbf{G}\mathbf{n}. \tag{4.2}$$

For simplicity, throughout this chapter we assume a ZF equaliser given by[1]

$$\mathbf{G} = \left(\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H, \tag{4.3}$$

and the output signal vector of the equaliser is expressed as

$$\mathbf{r} = \mathbf{Z}\mathbf{s} + \hat{\mathbf{n}}, \tag{4.4}$$

where $\hat{\mathbf{n}} = \mathbf{G}\mathbf{n}$ is the effective noise vector after equalisation. We note that the approaches detailed below can be straightforwardly extended to other equalisers, by considering the resulting effective channel after equalisation.

Based on [62, 64], the MC matrix with tunable load impedances can be derived and defined as (Appendix A)

$$\mathbf{Z} = \left[z_A \cdot \mathbf{I} + diag\left(\mathbf{z}_L\right)\right]\left[\mathbf{Z}_M + diag\left(\mathbf{z}_L\right)\right]^{-1}, \tag{4.5}$$

[1]We have extracted the mutual coupling matrix $\mathbf{Z}$ from the channel such that its effect can be exploited at the transmitter side by adaptive load tuning.

where $z_A$ denotes the fixed antenna load and $\mathbf{z}_L = \left[z_{L_1}, z_{L_2}, ..., z_{L_{N_t}}\right]^T$ is the tunable load impedance vector to be optimised. $\mathbf{Z}_M$ denotes the mutual impedance matrix that can be expressed as

$$\mathbf{Z}_M = \begin{bmatrix} z_A & z_{M_1} & z_{M_2} & \cdots & z_{M_{N_t-1}} \\ z_{M_1} & z_A & z_{M_1} & \ddots & \vdots \\ z_{M_2} & z_{M_1} & \ddots & \ddots & z_{M_2} \\ \vdots & \ddots & \ddots & \ddots & z_{M_1} \\ z_{M_{N_t-1}} & \cdots & z_{M_2} & z_{M_1} & z_A \end{bmatrix}. \tag{4.6}$$

In (4.6), each $z_{M_i}$ denotes the mutual impedance of two antenna elements with the distance of $k \cdot d_t$, $k \in \{1, 2, ..., N_t - 1\}$. The value of $z_A$ and $z_{M_i}$ can be obtained by the induced electromagnetic-field (EMF) method based on the normalised antenna spacing $d_t$ of the antenna array [63] (Appendix B).

## 4.3 Proposed Method based on CI Formulation

In this section, we firstly prove that the full elimination of the coupling effect is not achievable by solely tuning the values of the antenna loads, followed by the proposed methods that exploit the coupling effect based on the CI formulation via convex optimisation. For notational simplicity, we first rewrite the MC matrix in (4.5) as

$$\begin{aligned} \mathbf{Z} &= \left[z_A \cdot \mathbf{I} + diag\left(\mathbf{z}_L\right)\right]\left[z_A \cdot \mathbf{I} + diag\left(\mathbf{z}_L\right) + \left(\mathbf{Z}_M - z_A \cdot \mathbf{I}\right)\right]^{-1} \\ &= \mathbf{Z}_d\left(\mathbf{Z}_d + \tilde{\mathbf{Z}}_M\right)^{-1}, \end{aligned} \tag{4.7}$$

where $\mathbf{Z}_d$ is a diagonal matrix and can be expressed

$$\mathbf{Z}_d = z_A \cdot \mathbf{I} + diag\left(\mathbf{z}_L\right) = diag\left(z_A \cdot \mathbf{1} + \mathbf{z}_L\right), \tag{4.8}$$

where $\mathbf{1} = [1, 1, \cdots, 1]^T$. $\tilde{\mathbf{Z}}_M$ is obtained from $\mathbf{Z}_M$ with all the diagonal elements being zero, given by

$$\tilde{\mathbf{Z}}_M = \mathbf{Z}_M - z_A \cdot \mathbf{I}. \tag{4.9}$$

Before being transmitted through the wireless channel, the signal vector is affected by the coupling effect, which is fully represented by the multiplication of the MC matrix, as shown in (4.1) and (4.4). Therefore, we can regard the MC matrix as an interference matrix that generates the undesired inter-symbol interference. We note that in the ideal case without any coupling, $\mathbf{Z} = \mathbf{I}$ and no inter-symbol interference is observed. In the presence of the coupling effect, the equivalent signal vector to be transmitted becomes

$$\tilde{\mathbf{s}} = \mathbf{Z}\mathbf{s} = \mathbf{Z}_d \left(\mathbf{Z}_d + \tilde{\mathbf{Z}}_M\right)^{-1} \mathbf{s}. \tag{4.10}$$

Based on the knowledge of the linear algebra theory, the linear scaling of a vector can be equivalently represented by the vector multiplied by a diagonal matrix, and therefore $\tilde{\mathbf{s}}$ can also be expressed as

$$\tilde{\mathbf{s}} = diag\left(\mathbf{\Lambda}\right) \mathbf{s}, \tag{4.11}$$

where we introduce $\mathbf{\Lambda} = [\lambda_1, \lambda_2, \cdots, \lambda_{N_t}]^T$ and each data-dependent $\lambda_k = \frac{[\tilde{\mathbf{s}}]_k}{s_k}$ represents a scaling for the corresponding data symbol $s_k$. The property of $\mathbf{\Lambda}$ then fully characterises the coupling effect on the transmitted signals. It should be noted that the mathematical representation of (4.11) does not necessarily mean that there is no inter-symbol interference, because each $\lambda_k$ can be a complex number, which geometrically represents the existence of a phase rotation. Based on (4.10) and (4.11), we can obtain

$$\mathbf{Z}_d \left(\mathbf{Z}_d + \tilde{\mathbf{Z}}_M\right)^{-1} \mathbf{s} = diag\left(\mathbf{\Lambda}\right) \mathbf{s}. \tag{4.12}$$

(4.12) can be further transformed into

$$\begin{aligned} & \mathbf{s} = \left(\mathbf{Z}_d + \tilde{\mathbf{Z}}_M\right)\mathbf{Z}_d^{-1} diag\left(\mathbf{\Lambda}\right)\mathbf{s} \\ \Rightarrow\ & \mathbf{s} = diag\left(\mathbf{\Lambda}\right)\mathbf{s} + \tilde{\mathbf{Z}}_M\mathbf{Z}_d^{-1} diag\left(\mathbf{\Lambda}\right)\mathbf{s} \\ \Rightarrow\ & \left[diag\left(\mathbf{\Lambda}\right) - \mathbf{I}\right]\mathbf{s} + \tilde{\mathbf{Z}}_M\mathbf{Z}_d^{-1} diag\left(\mathbf{\Lambda}\right)\mathbf{s} = \mathbf{0}, \end{aligned} \tag{4.13}$$

which reveals the relationship between the scaling vector $\mathbf{\Lambda}$ and the load impedance matrix $\mathbf{Z}_d$. Accordingly, based on (4.13) we first consider the full elimination of MC, and further explore two different design approaches for $\mathbf{Z}_d$ based on CI.

### 4.3.1 Full Elimination of the MC Effect

We firstly consider the signal processing technique that fully eliminates the coupling effect. Mathematically, this is equivalent to $\tilde{\mathbf{s}} = \mathbf{s}$, which further leads to

$$\lambda_k = 1,\ \forall k \in \mathscr{N}, \tag{4.14}$$

where we denote $\mathscr{N} = \{1, 2, \cdots, N_t\}$. By substituting (4.14) into (4.13), the following condition should be satisfied

$$\tilde{\mathbf{Z}}_M\mathbf{Z}_d^{-1}\mathbf{s} = \mathbf{0}, \tag{4.15}$$

which means that for a specific transmit symbol vector $\mathbf{s}$, if we can adjust each value of the load impedance $z_{L_k}$ to satisfy (4.15), the effect of MC can be fully eliminated. Noting that $\mathbf{Z}_d$ is a diagonal matrix and $\mathbf{s}$ is a column vector, (4.15) can be further transformed into

$$\tilde{\mathbf{Z}}_M diag\left(\mathbf{s}\right) \mathrm{vec}\left(\mathbf{Z}_d^{-1}\right) = \mathbf{0}, \tag{4.16}$$

which leads to the following proposition.

**Proposition:** It is not achievable to fully eliminate the coupling effect by changing the values of each antenna load $z_{L_k}$, for data symbols from a normalised PSK or QAM constellation where $s_k \neq 0$, $\forall k$.

**Proof:** (4.16) can be regarded as a feasibility problem, where we need to obtain the value of each load impedance to satisfy this equation for a specific transmit symbol vector $\mathbf{s}$. This is equivalent to finding non-zero solutions for a linear system $\mathbf{Ax} = \mathbf{0}$. For this problem, $\mathbf{x} = \text{vec}\left(\mathbf{Z}_d^{-1}\right)$ is a column vector and $\mathbf{A} = \tilde{\mathbf{Z}}_M diag(\mathbf{s})$. Based on the linear algebra theory, the condition for such a linear system to have a non-zero solution is $\det\{\mathbf{A}\} = 0$, which means

$$\det\left\{\tilde{\mathbf{Z}}_M diag(\mathbf{s})\right\} = 0. \tag{4.17}$$

(4.17) is further transformed into the following condition

$$\det\left\{\tilde{\mathbf{Z}}_M\right\} = 0 \quad \text{or} \quad \det\{diag(\mathbf{s})\} = 0. \tag{4.18}$$

Since $diag(\mathbf{s})$ is a diagonal matrix and it is easy to observe from (4.9) that the determinant of $\tilde{\mathbf{Z}}_M$ is also not zero, it is therefore concluded that (4.16) cannot have non-zero solutions.

**Remark:** Based on the above proposition, (4.16) can only have zero solutions. With the definition of $\mathbf{Z}_d$ in (4.8), this means that the values of each $z_{L_k}$ will be extremely large in practice, such that $\mathbf{Z}_d^{-1}$ approaches $\mathbf{0}$. In this case, $z_A$ and each $z_{M_i}$ will be very small compared to $z_{L_k}$ and therefore can be negligible. Subsequently, (4.7) will be transformed into

$$\begin{aligned}\mathbf{Z} &= [z_A \cdot \mathbf{I} + diag(\mathbf{z}_L)]\,[\mathbf{Z}_M + diag(\mathbf{z}_L)]^{-1} \\ &\approx diag(\mathbf{z}_L)\,[diag(\mathbf{z}_L)]^{-1} \\ &= \mathbf{I}.\end{aligned} \tag{4.19}$$

However, this case is not feasible in practice, as an extremely large value of the load impedance will result in an extremely small current in the transmitting antenna, which leads to a very small power transfer efficiency [107, 108].

### 4.3.2 MC Exploitation based on CI

While the full elimination of the coupling effect is not achievable, in the following we show that this effect can be exploited to further benefit the MIMO detection performance based on the CI formulation in Chapter 2.2.4. Specific to the MC exploitation in this chapter, the constructive region relaxes the strict angle requirement for manipulating the interfering signals, and more importantly relaxes the value range of each load impedance, which makes it possible to obtain a constructive coupling matrix. Accordingly, different from the case in Chapter 4.3.1 where the full elimination is not possible, the exploitation of the coupling effect based on CI can be achieved with practical antenna load values. For a practical antenna array, the real part of each $z_{L_k}$ should be non-negative [109, 110, 111]. In the following, we employ optimisation approaches to determine the scaling vector under realistic implementation constraints, where we consider two optimisation criteria, as detailed below.

***Max-Min***: Recall (4.13), we first consider a max-min optimisation problem where the minimum value of $\Re(\lambda_k)$ is maximised, formulated based on the CI formulation in Chapter 2.2.4 as

$$\begin{aligned}
\mathscr{P}_{4.1}: \ & \max_{\lambda_k, z_{L_k}} \min_{k} \ \Re(\lambda_k) \\
s.t. \quad & [diag(\mathbf{\Lambda}) - \mathbf{I}]\,\mathbf{s} + \tilde{\mathbf{Z}}_M \mathbf{Z}_d^{-1} diag(\mathbf{\Lambda})\,\mathbf{s} = \mathbf{0} \\
& [\Re(\lambda_k) - 1] \cdot \tan(\theta_t) \geq |\Im(\lambda_k)|, \ \forall k \in \mathscr{N} \\
& \Re(z_{L_k}) \geq 0, \ \forall k \in \mathscr{N} \\
& \Re(\lambda_k) \geq 1, \ \forall k \in \mathscr{N}
\end{aligned} \tag{4.20}$$

By introducing $t_k$, $x_k$, $\mathbf{t}$ and $\mathbf{x}$ that are given by

$$\mathbf{t} = [t_1, t_2, \cdots, t_{N_t}]^T, \ \mathbf{x} = [x_1, x_2, \cdots, x_{N_t}]^T, \ t_k = \lambda_k - 1, \ x_k = \frac{\lambda_k}{z_A + z_{L_k}}, \ \forall k \in \mathscr{N}, \tag{4.21}$$

the max-min problem $\mathscr{P}_{4.1}$ in (4.20) can be further transformed into

$$\begin{aligned}
\mathscr{P}_{4.2}: \ & \max_{x_k, t_k} \ \omega \\
s.t. \quad & diag(\mathbf{t})\,\mathbf{s} + \tilde{\mathbf{Z}}_M diag(\mathbf{x})\,\mathbf{s} = \mathbf{0} \\
& \Re(t_k) \cdot \tan(\theta_t) \geq |\Im(t_k)|, \ \forall k \in \mathscr{N} \\
& \Re(z_{L_k}) \geq 0, \ \forall k \in \mathscr{N} \\
& \omega \leq \Re(t_k) + 1, \ \forall k \in \mathscr{N}
\end{aligned} \tag{4.22}$$

where $\omega$ is an introduced auxiliary variable. Based on (4.21), we can express the load impedance $z_{L_k}$ as

$$z_{L_k} = \frac{t_k + 1}{x_k} - z_A. \tag{4.23}$$

As each $t_k$ is a complex variable, the constraint $\Re(z_{L_k}) \geq 0$ is equivalent to

$$\begin{aligned}
& \Re\left(\frac{t_k + 1}{x_k}\right) \geq \Re(z_A) \\
\Rightarrow \ & [\Re(t_k) + 1]\,\Re(x_k) + \Im(t_k)\,\Im(x_k) \geq \Re(z_A) \cdot |x_k|^2.
\end{aligned} \tag{4.24}$$

Note that (4.24) is a complicated non-convex constraint that makes the optimisation non-convex. In the following, we perform a transformation such that this non-convex constraint can be transformed into a convex one, and the optimisation can be efficiently solved. To be more specific, by introducing a parameter $\mu_{th}$ multiplied to the right-hand side of the constraint, we remove the imaginary part multiplication $\Im(t_i)\,\Im(x_i)$ at the left-hand side, and (4.24) is transformed into

$$\begin{aligned}
& [\Re(t_k) + 1]\,\Re(x_k) + \Im(t_k)\,\Im(x_k) \geq \Re(z_A) \cdot |x_k|^2 \\
\Rightarrow & [\Re(t_k) + 1]\,\Re(x_k) \geq \mu_{th} \Re(z_A) \cdot |x_k|^2 \\
\Rightarrow & \Re(x_k) \geq \frac{\mu_{th} \Re(z_A)}{\Re(t_k) + 1} \cdot |x_k|^2,
\end{aligned} \tag{4.25}$$

where $\mu_{th} \geq 1$. With a larger value of $\mu_{th}$, there is a higher possibility that the optimisation problem returns a practical solution. However, a larger value of $\mu_{th}$ will also lead to a loss in the performance gains, as it reduces the feasible domain of the variables. This effect will be numerically studied by simulation results in Chapter 4.6. We further impose a relatively strong convex constraint to substitute the non-convex constraint (4.25), which is an upper-bound, such that the solution obtained with the upper-bound constraint also satisfies the original constraint. Particularly, based on $[\Re(t_k)+1] \geq 1$, if the following condition is satisfied

$$\Re(x_k) \geq \mu_{th}\Re(z_A)\cdot|x_k|^2, \tag{4.26}$$

the constraint in (4.25) is automatically satisfied. With this approach, the non-convex constraint (4.24) has been substituted with a convex constraint as shown in (4.26), and the final optimisation problem in a convex form is given by

$$\begin{aligned}
\mathscr{P}_{4.3}: \ & \max_{x_k,t_k} \ \omega \\
s.t. \quad & diag(\mathbf{t})\mathbf{s} + \tilde{\mathbf{Z}}_M diag(\mathbf{x})\mathbf{s} = \mathbf{0} \\
& \Re(t_k)\cdot\tan(\theta_t) \geq |\Im(t_k)|, \ \forall k \in \mathscr{N} \\
& \Re(x_k) \geq \mu_{th}\Re(z_A)\cdot|x_k|^2, \ \forall k \in \mathscr{N} \\
& \omega \leq \Re(t_k)+1, \ \forall k \in \mathscr{N}
\end{aligned} \tag{4.27}$$

$\mathscr{P}_{4.3}$ is a second-order cone programming (SOCP) optimisation and can be efficiently solved by convex optimisation tools or interior-point methods.

***Sum-Max***: We also pursue an optimisation problem that targets at maximising the average performance improvement for the data symbol. This

leads to a sum-max problem, constructed as

$$\begin{aligned}
\mathscr{P}_{4.4}: \ & \max_{\lambda_k, z_{L_k}} \ \frac{\sum_{k=1}^{N_t} \Re(\lambda_k)}{N_t} \\
s.t. \quad & [diag(\mathbf{\Lambda}) - \mathbf{I}]\mathbf{s} + \tilde{\mathbf{Z}}_M \mathbf{Z}_d^{-1} diag(\mathbf{\Lambda})\mathbf{s} = \mathbf{0} \\
& \Re(t_k) \cdot \tan(\theta_t) \geq |\Im(\lambda_k)|, \ \forall k \in \mathscr{N} \\
& \Re(z_{L_k}) \geq 0, \ \forall k \in \mathscr{N} \\
& \Re(\lambda_k) \geq 1, \ \forall k \in \mathscr{N}
\end{aligned} \tag{4.28}$$

which can be similarly transformed into a convex version, given by

$$\begin{aligned}
\mathscr{P}_{4.5}: \ & \max_{x_k, t_k} \ \sum_{k=1}^{N_t} \Re(t_k) \\
s.t. \quad & diag(\mathbf{t})\mathbf{s} + \tilde{\mathbf{Z}}_M diag(\mathbf{x})\mathbf{s} = \mathbf{0} \\
& \Re(t_k) \cdot \tan(\theta_t) \geq |\Im(t_k)|, \ \forall k \in \mathscr{N} \\
& \Re(x_k) \geq \mu_{th} \Re(z_A) \cdot |x_k|^2, \ \forall k \in \mathscr{N}
\end{aligned} \tag{4.29}$$

After the optimisation problems are solved, the values of each antenna load can be calculated based on (4.23). Based on the description of the CI formulation in Chapter 2.2.4, the extension to QAM modulations can be similarly obtained and is omitted for brevity.

## 4.4 Performance Metric

In this chapter, we present the performance metric for the proposed approaches, where we focus on the ZF equalisation method. The separately-correlated channel model employed is firstly characterised, followed by the derivation of the analytical probability of error.

### 4.4.1 Correlated Channel Characterisation

We firstly study the property of the separately-correlated channels employed in this chapter. Based on the channel modelling in Chapter 2.3.2 and note

that $\mathbf{H}_\alpha$ is diagonal, we can express the elements in $\mathbf{H}$ as

$$[\mathbf{H}]_{i,j} = \frac{1}{\sqrt{N_p}} \sum_{m=1}^{N_p} \alpha_m [\mathbf{A}_r]_{i,m} [\mathbf{A}_t]_{m,j}. \tag{4.30}$$

For ease of analysis, we only consider the expectation over the Rayleigh components $\mathbf{H}_\alpha$ and regard $\mathbf{A}_r$, $\mathbf{A}_t$ as constant matrices throughout the derivation. Noting that each $\alpha_m$ is independent and $\alpha_m \sim \mathbb{CN}(0,1)$, therefore each element in $\mathbf{H}$ is the summation of independent Gaussian variables, which follows a Gaussian distribution with zero mean. The resulting channel matrix $\mathbf{H}$ is therefore also Gaussian with zero mean. To calculate the probability of error, we need the knowledge of the diagonal terms of the covariance for the channel vector of each user $k$, and we consider the following statistics

$$\begin{aligned} \mathbf{C}_\mathbf{H} &= \mathbb{E}\left\{ [\mathbf{H} - \mathbb{E}(\mathbf{H})] \cdot [\mathbf{H} - \mathbb{E}(\mathbf{H})]^H \right\} \\ &= \mathbb{E}\left\{ \mathbf{A}_r \mathbf{H}_\alpha \mathbf{A}_t^H \mathbf{A}_t \mathbf{H}_\alpha^H \mathbf{A}_r^H \right\} \\ &= \mathbf{A}_r \cdot \mathbb{E}\left\{ \mathbf{H}_\alpha \mathbf{A}_t^H \mathbf{A}_t \mathbf{H}_\alpha^H \right\} \cdot \mathbf{A}_r^H \\ &= \mathbf{A}_r \cdot \mathbb{E}\{\mathbf{T}\} \cdot \mathbf{A}_r^H \end{aligned} \tag{4.31}$$

where for simplicity we denote

$$\mathbf{T} = \mathbf{H}_\alpha \mathbf{A}_t^H \mathbf{A}_t \mathbf{H}_\alpha^H = \mathbf{H}_\alpha \mathbf{A} \mathbf{H}_\alpha^H, \tag{4.32}$$

where $\mathbf{A} = \mathbf{A}_t^H \mathbf{A}_t$. Note that $\mathbf{H}_\alpha$ is diagonal, and each element in $\mathbf{T}$ can be calculated as

$$[\mathbf{T}]_{i,j} = \frac{1}{N_p} \alpha_i \alpha_j^H [\mathbf{A}]_{i,j}. \tag{4.33}$$

Based on the fact that each entry in $\mathbf{H}_\alpha$ is independent, we obtain

$$\begin{aligned} &\mathbb{E}\left\{ \alpha_i \alpha_j^H \right\} = 0, \ \forall i \neq j, \\ &\mathbb{E}\left\{ \alpha_i \alpha_i^H \right\} = 1, \ \forall i \in \left\{ 1, 2, \cdots, N_p \right\} \end{aligned} \tag{4.34}$$

which further leads to

$$\begin{aligned}\mathbb{E}\left\{[\mathbf{T}]_{i,j}\right\} &= 0,\ \forall i \neq j,\\ \mathbb{E}\left\{[\mathbf{T}]_{i,i}\right\} &= \frac{1}{N_p}\cdot\left[\mathbf{A}_{diag}\right]_{i,i},\end{aligned} \tag{4.35}$$

where $\mathbf{A}_{diag} = diag\left\{[\mathbf{A}]_{1,1}, [\mathbf{A}]_{2,2}, \cdots, [\mathbf{A}]_{N_p,N_p}\right\}$ is a diagonal matrix formed from the diagonal entries of $\mathbf{A}$. With the above derivation, (4.31) can finally be obtained as

$$\mathbf{C_H} = \frac{1}{N_p}\cdot\mathbf{A}_r\mathbf{A}_{diag}\mathbf{A}_r^H. \tag{4.36}$$

### 4.4.2 Probability of Error

Based on the above results, we now study the analytical probability of error for the proposed approaches. Specifically, for conventional ZF receiver with PSK modulations, the analytical BER over Rayleigh fading channels is given by (2.7) [30, 112, 113]. For the proposed method where the coupling effect is exploited to benefit the system performance, each transmit symbol $s_k$ is enhanced by a factor $\lambda_k$, which can be translated as an increase in the transmit power, and $|\lambda_k|^2$ can represent the power improvement factor, without incurring additional interference. Therefore, the analytical BER for the proposed approaches based on the ZF equaliser can also be obtained by (2.7), where $\hat{\gamma}_k$ for the proposed scheme is obtained as

$$\hat{\gamma}_k = \frac{|\lambda_k|^2}{\log_2(\mathbb{N})\cdot\sigma^2\left[\mathbf{C}_\mathbf{H}^{-1}\right]_{k,k}}. \tag{4.37}$$

## 4.5 Practical Implementation Aspects

In this section, the practical implementation of the proposed techniques is explored, where a number of practical implementation issues are discussed.

### 4.5.1 Extraction of H

From (4.3), we observe that the receiver needs to obtain **H** to perform the equalisation, while we note that we can only obtain **HZ** using generic chan-

nel estimation techniques with pilots. Therefore, we need to extract $\mathbf{H}$ from $\mathbf{HZ}$. We note that the mutual impedance matrix $\mathbf{Z}_M$ is only dependent on the array structure and does not change, and accordingly $\mathbf{Z}_M$ is known to the transmitter either by the induced EMF method or other experimental measurements, for example the boundary-value approach or the transmission-line method [63]. At the pilot stage, the channel is obtained by first setting the value of each antenna load to a specific reference value, for example $z_{L_k} = 50\Omega, \forall k \in \mathscr{N}$, which is assumed known to both the transmitter and the receiver. We denote the resulting load impedance vector as $\mathbf{z}_L^0$ and MC matrix as $\mathbf{Z}_0$. With standard channel estimation techniques, we can obtain $\mathbf{HZ}_0$ and further extract $\mathbf{H}$ from $\mathbf{HZ}_0$ to obtain the equaliser $\mathbf{G}$, as $\mathbf{Z}_0$ is already known to the receiver. Accordingly, the proposed approaches that exploit the constructive coupling effect can be employed during data transmission.

### 4.5.2 Use of Look-Up Tables

Based on (4.13) and the description of the proposed techniques, we observe that the optimal values of the analog loads are solely dependent on the transmit symbol vector, while irrespective of the channel matrix. This enables the design of a look-up table, in which the optimal values of the antenna loads can be directly obtained based on the transmit symbol vector. With this approach, the optimisation process can be conducted off-line to obtain the optimal values of the load impedances for each possible data symbol combination, prior to data transmission. This information can then be kept and used for future data transmission, and a symbol-level optimisation is therefore no longer needed, which significantly reduces the computational costs. If we assume $N$-PSK modulation is employed, the total number of possible combinations for the transmit symbol vector will be $L = N^{N_t}$, which is equal to the length of the look-up table. To illustrate this approach intuitively, an example of the look-up table is shown in Table 4.1, where we assume a $2 \times 2$ MIMO system with QPSK modulation, leading to a total number of 16 possible combinations of the transmit symbol vector. Without loss of generality,

we denote the 4 constellation points of QPSK as

$$m_1 = \frac{1}{\sqrt{2}} + j \cdot \frac{1}{\sqrt{2}},\ m_2 = \frac{1}{\sqrt{2}} - j \cdot \frac{1}{\sqrt{2}},\ m_3 = -\frac{1}{\sqrt{2}} - j \cdot \frac{1}{\sqrt{2}},\ m_4 = -\frac{1}{\sqrt{2}} + j \cdot \frac{1}{\sqrt{2}}. \tag{4.38}$$

In the look-up table, the optimal load values are obtained based on the optimisation $\mathscr{P}_{4.3}$. In Table 4.1, $d_t$ and $d_r$ represents the equidistant antenna spacing normalised by the carrier wavelength at the transmitter and receiver, respectively. $d_l$ denotes the normalised dipole length of the antenna array.

| Transmit Symbol Vector | | Optimal Impedance Values (Ω) | |
|---|---|---|---|
| Antenna 1 | Antenna 2 | Antenna 1 | Antenna 2 |
| $m_1$ | $m_1$ | 50 | 50 |
| $m_1$ | $m_2$ | 43.546-40.203$j$ | 52.009+26.114$j$ |
| $m_1$ | $m_3$ | 45.812-7.396$j$ | 45.812-7.396$j$ |
| $m_1$ | $m_4$ | 52.053+26.096$j$ | 43.532-40.181$j$ |
| $m_2$ | $m_1$ | 49.217+27.758$j$ | 44.265-41.923$j$ |
| $m_2$ | $m_2$ | 50 | 50 |
| $m_2$ | $m_3$ | 43.567-40.277$j$ | 51.902+26.201$j$ |
| $m_2$ | $m_4$ | 45.812-7.396$j$ | 45.812-7.396$j$ |
| $m_3$ | $m_1$ | 45.812-7.396$j$ | 45.812-7.396$j$ |
| $m_3$ | $m_2$ | 52.053+26.096$j$ | 43.532-40.181$j$ |
| $m_3$ | $m_3$ | 50 | 50 |
| $m_3$ | $m_4$ | 43.546-40.203$j$ | 52.009+26.114$j$ |
| $m_4$ | $m_1$ | 43.567-40.277$j$ | 51.902+26.201$j$ |
| $m_4$ | $m_2$ | 45.812-7.396$j$ | 45.812-7.396$j$ |
| $m_4$ | $m_3$ | 49.217+27.758$j$ | 44.265-41.923$j$ |
| $m_4$ | $m_4$ | 50 | 50 |

**Table 4.1:** Look-up Table for $2 \times 2$ MIMO, QPSK, $\mathscr{P}_{4.3}$, $d_t$=0.2, $d_r$=0.5, $d_l$=0.3, $\mu_{th}$=1.1

### 4.5.3 Solutions for Extreme Load Values

It should be noted that, for some transmit symbol combinations, the solutions returned by convex optimisation lead to extremely large negative values of load impedances, which is unreasonable and means that there exist no practical solutions for such symbol combinations.[2] In the example of Table 4.1, this happens when identical symbols are transmitted, which are $[m_1, m_1]^T$, $[m_2, m_2]^T$, $[m_3, m_3]^T$ and $[m_4, m_4]^T$.

[2]We note that the value range of the imaginary part of a practical tunable load is only hundreds of Ω.

Intuitively, this phenomenon is extended to larger transmit antenna numbers, where no practical solutions can be obtained when all the antennas transmit the same symbols. In this case, the fixed coupling matrix is used and the value for all the antenna loads is $z_{L_k} = 50\Omega$, $\forall k \in \mathscr{N}$, as shown in Table 4.1. Furthermore, based on the above analysis, we can obtain the feasibility probability of the proposed approaches for $\mathbb{N}$-PSK, given by

$$P_{feasibility} = 1 - \frac{\mathbb{N}}{\mathbb{N}^{N_t}} = 1 - \frac{1}{\mathbb{N}^{N_t-1}}. \tag{4.39}$$

We further note the value of $\mu_{th}$ also has an impact on the feasibility, and therefore $P_{feasibility}$ in (4.39) can be regarded as an upper bound. The effect of $\mu_{th}$ on the feasibility probability will be numerically studied in the simulation results.

### 4.5.4 Symbol-by-Symbol Switching of Loads

Note that the proposed optimisations necessitate the adaptation of the load impedance values $z_{L_k}$ on a symbol level, where for each data transmission, the value of each load impedance is tuned based on the look-up table to enable the application of the proposed methods. For the fast tuning of the analog loads, it has been shown in [114] and the references therein that varactor technologies that support adaptive impedance tuning can be divided into 3 categories: semiconductor-based varactor diodes, microelectromechanical-system varactors, and ferroelectric-based varactors. Specifically, semiconductor-based varactor diodes and ferroelectric-based varactors can support the tuning speed as fast as 100ns-100$\mu$s [114, 115, 116].[3] Semiconductor-based varactor diodes are more attractive in low power design, while ferroelectric-based varactors are capable for high power design. The tuning of the load impedances can be realised by a matching system based on an automated impedance tuning unit with ferroelectric varactors. Furthermore, recent studies on ESPARs, where the radiation patterns

[3]For example, the symbol rate in LTE is 66.7$\mu$s, which means that the symbol-by-symbol switching of the loads can be supported.

are formed by tuning the load impedance for each parasitic element, have shown that the frequent tuning of the load impedance is achievable [8, 14, 15, 21, 22]. This is verified by the proof-of-concept experiments in [117], which supports the implementation of the proposed MC exploitation in practical wireless networks. The study of the downlink precoding techniques for ESPARs will be presented in Chapter 6.

## 4.6 Numerical Results

In this section, we evaluate the usefulness of the proposed approaches based on Monte Carlo simulations. We assume a carrier frequency of $f$=2.6GHz, and the separately-correlated channels are generated based on Chapter 2.3.2 with $N_p = 50$. A dipole antenna array with a normalised dipole length of $d_l = 0.3$ is assumed, where the normalised antenna spacing at the transmitter and receiver is $d_t = 0.2$ and $d_r = 0.5$, respectively. In the simulations we consider a $4 \times 4$ P2P MIMO system. For simplicity, the ZF receiver is employed throughout the simulations, while the performance benefits extend to other receiver structures. The above parameters remain constant throughout the simulations unless otherwise stated. For clarity, the following abbreviations are employed throughout the simulations.

- 'ZF no MC': ZF receiver under the ideal case with no MC effect, and this case is for reference only;
- 'ZF with MC': ZF receiver with fixed coupling effect, where the value of each antenna load is 50Ω;
- 'max-min': the proposed approach based on the 'max-min' optimisation problem $\mathscr{P}_{4.3}$;
- 'sum-max': the proposed approach based on the 'sum-max' optimisation problem $\mathscr{P}_{4.5}$.

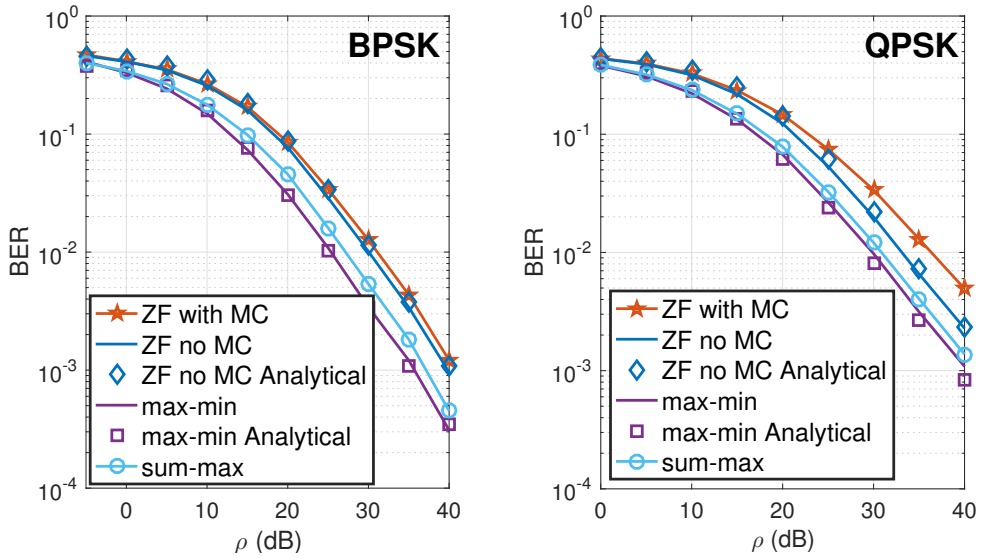
Fig. 4.2 presents the BER performance of the proposed approaches for BPSK and QPSK with respect to the increasing transmit SNR. For both modulations, it is observed that 'ZF with MC' is inferior to 'ZF no MC' and

achieves the worst BER performance, due to the detrimental coupling effect. For the proposed methods that exploit the MC effect, it can be observed that both the 'max-min' and 'sum-max' approach outperform 'ZF no MC' and 'ZF with MC'. Specifically, an SNR gain of 5dB for BPSK and 7dB for QPSK is observed compared to 'ZF with MC', as the antenna coupling effect is exploited to improve the performance. Moreover, for both modulations it is observed that 'max-min' outperforms 'sum-max' and achieves the best BER performance. As for the analytical results, a close match can be observed, which validates our analyses in Chapter 4.4.

In Fig. 4.3, we evaluate the rate performance of the proposed approaches in terms of the average throughput, where the throughput is defined as

$$\begin{aligned} T &= (1-\text{BLER}) \cdot m \cdot N_t \ bits/channel\ use \\ &= (1-P_b)^{F \cdot m} \cdot m \cdot N_t \ bits/channel\ use. \end{aligned} \tag{4.40}$$

In (4.40), BLER denotes the block error rate, and $P_b$ is the BER. $m$ denotes the information bits per symbol, and the block length used for the simulations




**(a)** BPSK **(b)** QPSK

**Figure 4.2:** Analytical and simulated BER v.s. transmit SNR, $4 \times 4$, $d_t = 0.2$, $d_r = 0.5$, $d_l = 0.3$, $\mu_{th} = 1.1$

is $F = 20$ symbols. For both modulations, it can be observed that 'ZF with MC' achieves the lowest throughput performance due to a higher BER. For both proposed methods, we observe a throughput benefit compared to the conventional case with and without MC, as the coupling effect is further exploited. It is not surprising to observe that the method based on 'max-min' also outperforms 'sum-max' in terms of the average throughput.

In Fig. 4.4, the BER performance of the proposed schemes for 16QAM is shown with respect to the increasing transmit SNR. It is first observed that the coupling effect has a much more severe impact on the BER performance for 16QAM. For the proposed approaches, a similar trend can be observed where the proposed methods can still offer performance gains. The 'max-min' and 'sum-max' approaches achieve a similar BER performance for 16QAM, and we can observe an around 2dB gain compared to the ideal ZF case without MC. We should note that 'ZF no MC' here is only shown as a reference, and the performance gain over the practical 'ZF with MC' is still large at the high SNR regime, which means that the proposed approaches are also advantageous for QAM modulations.

Fig. 4.5 compares the BER performance of each scheme with respect to

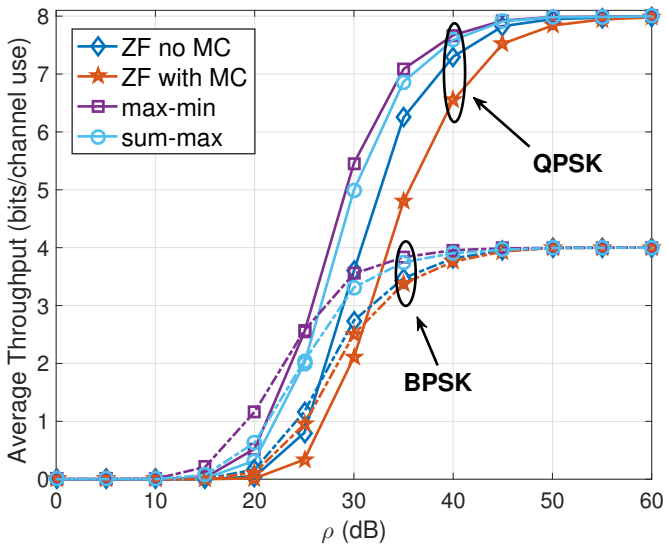



**Figure 4.3:** Average throughput v.s. transmit SNR, $4 \times 4$, $d_t = 0.2$, $d_r = 0.5$, $d_l = 0.3$, $\mu_{th} = 1.1$, $F = 20$ symbols per frame, BPSK and QPSK

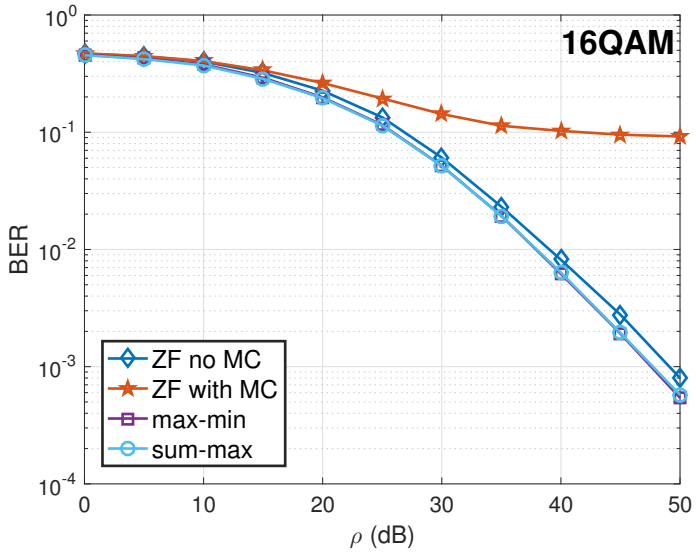
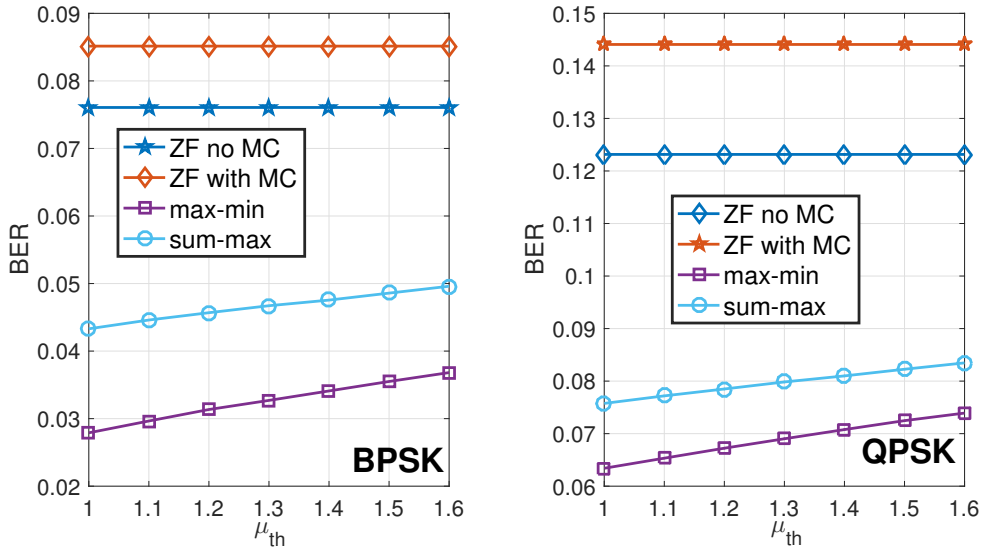



**Figure 4.4:** BER v.s. transmit SNR, $4 \times 4$, $d_t = 0.2$, $d_r = 0.5$, $d_l = 0.3$, 16QAM




**(a)** BPSK **(b)** QPSK

**Figure 4.5:** BER v.s. value of $\mu_{th}$, $4 \times 4$, $d_t = 0.2$, $d_r = 0.5$, $d_l = 0.3$, SNR=20dB

the optimisation parameter $\mu_{th}$ for both BPSK and QPSK. For the proposed MC exploitation approaches based on CI, it is seen that the BER becomes slightly worse with an increase in the value of $\mu_{th}$, because a larger value of $\mu_{th}$ reduces the feasible region of the antenna loads, as seen in $\mathscr{P}_{4.3}$ and $\mathscr{P}_{4.5}$. We also observe a consistent performance gain for 'max-min' compared to

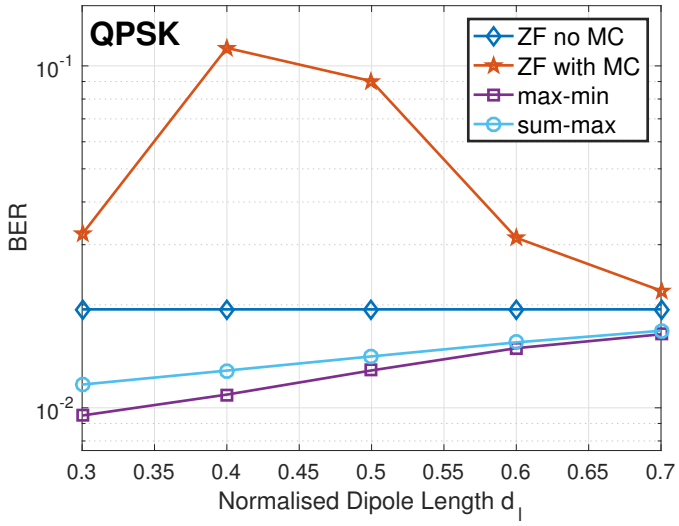



**Figure 4.6:** BER performance v.s. normalized dipole length, $4 \times 4$, $d_t = 0.2$, $d_r = 0.5$, $\mu_{th} = 1.1$, SNR=30dB, QPSK

'sum-max'.

In the following simulations, without loss of generality we focus on the QPSK modulation. In Fig. 4.6, the BER performance is compared with respect to the normalised dipole length. For 'ZF with MC', the most detrimental coupling effect is observed when the dipole length is $d_l = 0.4$, which greatly degrades the system performance. 'ZF no MC' does not include any coupling effect and therefore its performance remain constant with the increasing dipole length. For the proposed approaches, it is observed that they outperform both 'ZF with MC' and 'ZF no MC' with the exploitation of MC. Generally, techniques based on 'max-min' are shown to achieve the best performance for all values of the dipole length. With the increase in the dipole length, the BER performance becomes slightly worse. This is because the value of $\Re(z_A)$ becomes larger when the dipole length increases (Appendix B), which reduces the feasible region of each tunable load and therefore leads to a worse BER performance.

Fig. 4.7 presents the BER with an increase in the antenna spacing at the transmitter. When the antenna spacing is small, 'ZF with MC' achieves a much worse BER performance compared to other schemes due to the strong

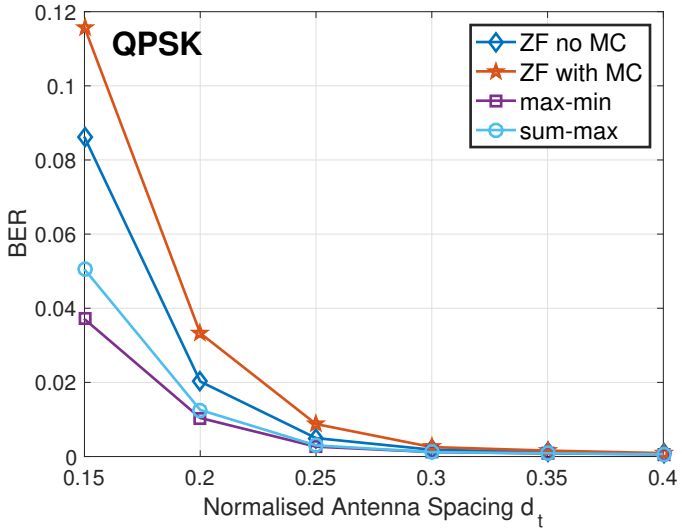



**Figure 4.7:** BER v.s. normalised antenna spacing at the transmitter, $4 \times 4$, $d_r = 0.5$, $d_l = 0.3$, $\mu_{th} = 1.1$, SNR=30dB, QPSK

coupling and correlation effect among the antenna array. For the proposed techniques, it is observed that they outperform both 'ZF with MC' and 'ZF no MC' as the MC effect is exploited to further benefit the system performance. With the increase in the antenna spacing, the performance gain between the proposed schemes and 'ZF with MC' becomes smaller as the coupling effect becomes weak, while the BER performance of all approaches is improved due to the reduced correlation effect when the antenna spacing is increased. The proposed techniques that exploit the MC effect therefore enable the design of compact antenna arrays.

Fig. 4.8 depicts the BER performance with respect to the number of transmit antennas, where we assume $N_r = N_t$ at SNR=30dB. With an increase in the antenna number, while the BER performance becomes worse for each scheme, a consistent performance gain can still be observed for the proposed methods over conventional ZF receiver with a fixed coupling effect. The performance gain of the proposed schemes persists with the number of antennas increasing. Moreover, for all cases it is observed that the proposed approach based on 'max-min' criterion outperforms the 'sum-max'.

Fig. 4.9 illustrates the feasibility probability of the 'max-min' optimisa-

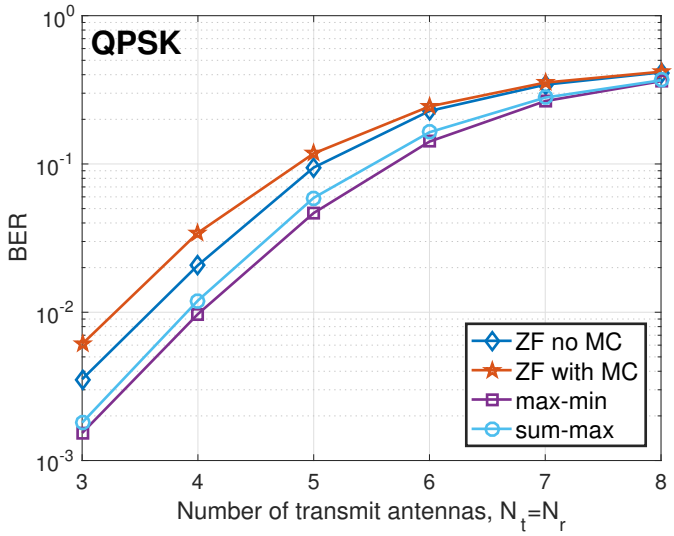



**Figure 4.8:** BER v.s. number of transmit antennas, $N_t = N_r$, $d_t = 0.2$, $d_r = 0.5$, $d_l = 0.3$, $\mu_{th} = 1.1$, SNR=30dB, QPSK

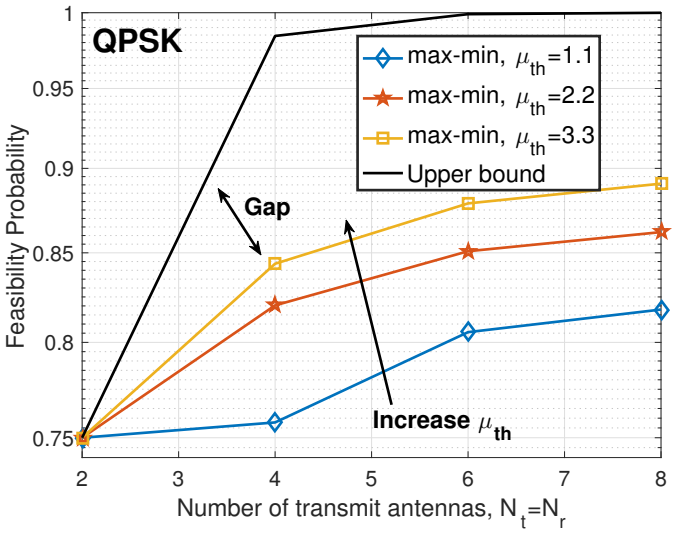



**Figure 4.9:** Feasibility probability v.s. antenna number, $N_r = N_t$, $d_t = 0.2$, $d_r = 0.5$, $d_l = 0.3$, QPSK

tion $\mathscr{P}_{4.3}$ with respect to the number of antennas, where the upper bound is obtained by (4.39). As can be seen, with an increase in the optimisation parameter $\mu_{th}$, the feasibility improves while the performance will be degraded, as evidenced by Fig. 4.5, which reveals the existence of a tradeoff between the performance and the feasibility.

## 4.7 Summary

In this chapter, we study the MC effect among antenna elements and propose techniques that further exploit this effect for P2P MIMO systems. While it is first shown that the full elimination of the coupling effect is not realisable via the tunable antenna loads, approaches that exploit the MC effect to further benefit the performance are proposed. The optimal values of each antenna load are obtained via convex optimisation methods based on the CI formulation, where practical constraints for the antenna array are considered. The implementation of the proposed schemes is also discussed, where a look-up table is introduced for practical application. The performance advantages of the proposed approaches over the cases with fixed coupling effect have been validated via extensive numerical results, especially in compact antenna deployments where the spacing between antenna elements is small.

# Chapter 5

# Mutual Coupling Exploitation for Multiuser Systems

This chapter is based on our works in [C5], [C9], [J3].

## 5.1 Introduction

In the previous chapter, we have proposed a MC exploitation technique for P2P MIMO systems. In this chapter, we extend our study on the MC exploitation to multi-user multiple-antenna systems.

Specific in this chapter, we propose a joint AD transmission approach, which exploits the coupling effect rather than eliminating this effect, to further improve the performance of multiple-antenna systems based on convex optimisation. In the proposed method, each antenna element is equipped with a tunable load to control the coupling effect and facilitate analog processing. By judiciously selecting the value of each antenna load through an iterative approach, the antenna coupling matrix and the precoding matrix are jointly optimised, which leads to an improved performance compared to existing approaches. Nevertheless, the joint optimisation results in a relatively high computational cost. For this reason, we further propose a decoupled low-complexity precoding approach to alleviate the computational burden of the joint scheme, where the digital precoding matrix is firstly obtained, based on which the optimisation on the tunable loads is performed.

The decoupled approach also enables the direct combination of the MC exploitation with existing closed-form digital precoding methods, where only an optimisation on the tunable antenna loads is needed. Moreover, since it may be infeasible to obtain any continuous load values in real systems, a more practical case with quantised antenna loads on the optimisation of the MC exploitation is considered, where a sequential search method based on the greedy algorithm is proposed.

Our study is further extended to the case of imperfect CSI, where we consider both the statistical CSI error model and the norm-bounded CSI error model, as introduced in Chapter 2.3.4. By calculating the SINR of the received symbols, we prove the optimality of the proposed approaches for closed-form precoding techniques in imperfect CSI scenarios. For optimisation-based precoding approaches, we further propose the robust precoding methods for these two CSI error scenarios, where the robust scheme is designed by incorporating the equivalent noise power for the statistical CSI errors, while the worst-case SINR is optimised for the norm-bounded CSI errors. Simulation results reveal the performance advantages of both the joint and the decoupled precoding approaches over the conventional digital precoding techniques with fixed MC effect, in both perfect CSI and imperfect CSI scenarios.

Accordingly, the profound conclusion from our works in Chapter 4 and Chapter 5 is that, hardware imperfections in the form of MC can be beneficial and further exploited to improve the system performance.

For clarity, the contributions of this chapter is summarised as

- We construct a joint AD precoding scheme that exploits the coupling effect for MU-MISO systems, where the optimal precoding and load values are obtained.
- We further propose a decoupled method to alleviate the high computational cost of the joint scheme. This decoupled method enables the use of closed-form precoders for the MC exploitation.

- For the practical case where the quantised load values are employed, we propose a sequential search approach based on a greedy algorithm to efficiently obtain the desired load values for each antenna element.

- We extend our study to imperfect CSI. By analysing the received SINR, we show that the proposed approaches can be directly extended to imperfect CSI for closed-form precoders, and the optimality still holds.

- For optimisation-based precoding approaches, we further propose the robust precoding methods for both the statistical CSI errors and the norm-bounded CSI errors.

## 5.2 System Model

We consider a MU-MISO system downlink, as shown in Fig. 5.1, where a BS with $N_t$ antennas each with its own RF chain and tunable load, communicates with $K$ single-antenna users simultaneously, and $K \leq N_t$. By employing a downlink precoding matrix $\mathbf{P} \in \mathbb{C}^{N_t \times K}$ at the BS, the received signal vector after the channel can be obtained as

$$\mathbf{y} = \mathbf{D}\mathbf{x} + \mathbf{n} = \mathbf{D}\mathbf{P}\mathbf{s} + \mathbf{n}, \tag{5.1}$$

where $\mathbf{n} = [n_1, n_2, ..., n_K]^T$ represents the additive Gaussian noise vector. $\mathbf{s} \in \mathbb{C}^{K \times 1}$ denotes the symbol vector, where for simplicity we assume that each element in $\mathbf{s}$ is from a normalised constellation and satisfies $\mathbb{E}\left\{\mathbf{s}\mathbf{s}^H\right\} = \mathbf{I}$. $\mathbf{D} \in \mathbb{C}^{K \times N_t}$ denotes the channel matrix, and in this chapter a geometric semi-correlated channel is considered, as described in Chapter 2.3.2. By incorporating the MC matrix in the channel to characterise the coupling effect, the channel vector for the $k$-th user is expressed as

$$\mathbf{d}_k = \mathbf{g}_k \mathbf{A}_k \mathbf{Z}, \tag{5.2}$$

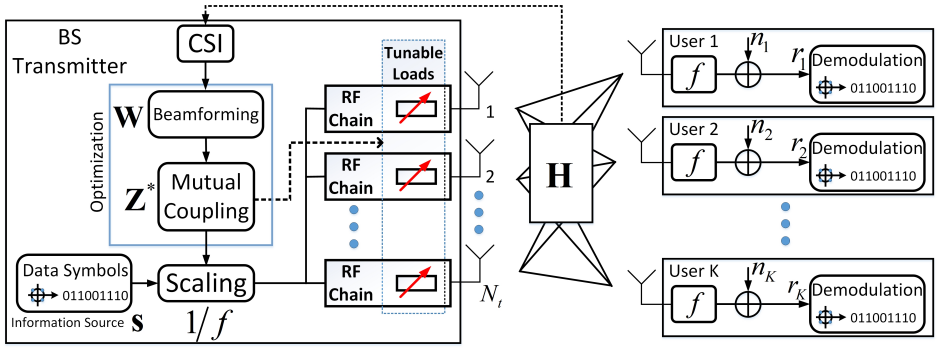



**Figure 5.1:** System model of a MU-MISO system with MC exploitation

where $\mathbf{D} = \left[\mathbf{d}_1^T, \mathbf{d}_2^T, ..., \mathbf{d}_K^T\right]^T$. $\mathbf{Z}$ represents the MC matrix as in Chapter 4, given by

$$\mathbf{Z} = \left[z_A \cdot \mathbf{I} + diag\left(\mathbf{z}_L\right)\right]\left[\mathbf{Z}_M + diag\left(\mathbf{z}_L\right)\right]^{-1}. \tag{5.3}$$

By decomposing $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_K]$, the received signal for the $k$-th user can be obtained based on (5.1) as

$$y_k = \mathbf{d}_k \mathbf{P}\mathbf{s} + n_k = \mathbf{d}_k \mathbf{p}_k s_k + \mathbf{d}_k \sum_{i \neq k}^{K} \mathbf{p}_i s_i + n_k. \tag{5.4}$$

Based on (5.4), we can express the received SINR of the $k$-th user as

$$\gamma_k = \frac{\left|\mathbf{d}_k \mathbf{p}_k\right|^2}{\sum\limits_{i \neq k}^{K} \left|\mathbf{d}_k \mathbf{p}_i\right|^2 + \sigma^2}, \tag{5.5}$$

where $\sigma^2$ is the noise power. In this chapter, perfect CSI is firstly assumed, while the effect of imperfect CSI is further investigated in the following.

## 5.3 Joint Analog-Digital Precoding Approach

In this section, we introduce the proposed joint AD precoding approach. We note that the joint optimisation can offer the best performance at the cost of a relatively high computational cost. Therefore, it is more suitable for small-scale MIMO systems in practice or as a performance benchmark.

With a tunable load employed for each antenna element, the mutual coupling effect can be controlled by modifying the value of each load, as

introduced in Chapter 4 in detail. For MU-MISO systems in this chapter, we propose to jointly optimise the precoding matrix in the digital domain and each antenna load in the analog domain, such that the resulting precoding matrix and load impedances are jointly optimal, which therefore improves the system performance. To exploit the MC effect, we firstly rewrite $\mathbf{D} = \mathbf{HZ}$, where $\mathbf{H}$ is given by

$$\mathbf{H} = \left[\mathbf{h}_1^T, ..., \mathbf{h}_k^T, ..., \mathbf{h}_K^T\right]^T, \tag{5.6}$$

and $\mathbf{h}_k = \mathbf{g}_k \mathbf{A}_k$ as in the semi-correlated channel model. With this decomposition, (5.1) can be rewritten as

$$\mathbf{y} = \mathbf{HZPs} + \mathbf{n}. \tag{5.7}$$

Based on (5.7), we construct the proposed precoding matrix $\mathbf{P}$ as

$$\mathbf{P} = \frac{1}{f} \cdot \mathbf{Z}^{-1}\mathbf{W}. \tag{5.8}$$

Following Chapter 2.2.1, $f$ is a scaling factor that ensures the signal power is not changed after precoding, given by

$$f = \left\|\mathbf{Z}^{-1}\mathbf{W}\right\|_F. \tag{5.9}$$

The introduction of the scaling factor $f$ for the constructed precoding matrix in (5.8) indeed contains an implicit power constraint and ensures that the transmit power is the same as the power of the data symbol vector, which is unlike the generic SINR balancing problem that contains a power budget in the optimisation constraints [40, 49]. As will be shown in the following, our construction of the optimisation $\mathscr{P}_{5.1}$ does not have such an explicit power constraint.

Accordingly, with the proposed precoding structure, (5.7) can be further transformed into

$$\mathbf{y} = \frac{1}{f} \cdot \mathbf{HWs} + \mathbf{n}. \tag{5.10}$$

As observed from (5.10), the proposed precoding structure fully eliminates the coupling effect in the channel, while the MC still has an effect on the system performance, which is fully characterised by the resulting scaling factor $f$. With the proposed precoding method, the received SINR for the $k$-th user in (5.5) is transformed into

$$\gamma_k = \frac{|\mathbf{h}_k \mathbf{w}_k|^2}{\sum\limits_{i \neq k}^{K} |\mathbf{h}_k \mathbf{w}_i|^2 + f^2 \sigma^2}, \tag{5.11}$$

where $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_K]$, and $f$ can accordingly be seen as a noise amplification factor. In (5.11), it is observed that the received SINR $\gamma_k$ is related to both $\mathbf{w}_k$ and $f$. By jointly optimising $\mathbf{W}$ and selecting the values for each load impedance, we maximise the minimum received SINR, which leads to the following optimisation problem

$$\begin{aligned} \mathscr{P}_{5.1}: \ & \max_{\mathbf{z}_L, \mathbf{W}} \min_{k} \ \gamma_k \\ s.t. \quad & \Re(z_{L_i}) \geq 0, \ \forall i \in \mathscr{N} \end{aligned} \tag{5.12}$$

where $\mathscr{N} = \{1, 2, \cdots, N_t\}$. The constraint on the value of each $z_{L_k}$ is from the fact that the real part of practical load impedances should be positive, such that the antenna array can radiate power [109, 110, 111]. Based on the expression of $f$, the optimisation problem $\mathscr{P}_{5.1}$ in (5.12) is not jointly convex in $\mathbf{z}_L$ and $\mathbf{W}$, and cannot be directly solved. Nevertheless, we note that $\mathscr{P}_{5.1}$ is a bi-convex optimisation problem, i.e., it is convex with respect to $\mathbf{W}$ when $\mathbf{z}_L$ is fixed and vice versa. Therefore, in the following we propose an alternating optimisation approach to obtain the optimal $\mathbf{W}$ and $\mathbf{z}_L$.

### 5.3.1 Solving W When $\mathbf{z}_L$ is Fixed

When $\mathbf{z}_L$ is fixed, the optimisation problem $\mathscr{P}_{5.1}$ is reduced to a typical FD precoding problem, which can be expressed as

$$\begin{aligned} \mathscr{P}_{5.2}: \ & \max_{\mathbf{W}} \min_{k} \ \gamma_k \\ s.t. \quad & \left\|\mathbf{Z}^{-1}\mathbf{W}\right\|_F^2 = p_0 \end{aligned} \tag{5.13}$$

**Remark on the Formulation of $\mathscr{P}_{5.2}$:** In $\mathscr{P}_{5.2}$, we note that the constraint is only a temporary power constraint for this FD SINR balancing sub-problem. Indeed, $M$ in $\mathscr{P}_{5.2}$ can be any arbitrary positive real value. This will not have an impact on the solution of the original optimisation problem $\mathscr{P}_{5.1}$, due to the existence of the scaling factor $f$ that constrains the power of the precoded signals. To be more specific, firstly we denote the solution of the optimisation problem $\mathscr{P}_{5.2}$ with the constraint value $p_0 = 1$ as $\mathbf{Z}^*$ and $\mathbf{w}_k^*$. Then, without loss of generality we consider the optimisation problem where $p_0 \neq 1$, and denote the solution as $\hat{\mathbf{Z}}$ and $\hat{\mathbf{w}}_k$. It is then easy to obtain that the solution for the reconstructed problem is $\hat{\mathbf{Z}} = \mathbf{Z}^*$ and $\hat{\mathbf{w}}_k = \sqrt{p_0} \cdot \mathbf{w}_k^*$. It is then easy to verify that these two problems return the same received SINR value. The above explanation therefore clarifies that the constraint $\left\|\mathbf{Z}^{-1}\mathbf{W}\right\|_F^2 = p_0$ is indeed only a temporary power constraint for the SINR balancing sub-problem and will not have an impact on the solution of the original problem.

Furthermore, it is shown in [40, 49] that $\mathscr{P}_{5.2}$ is equivalent to the following optimisation

$$\begin{aligned} \mathscr{P}_{5.3}: \ & \max_{\mathbf{W}} \min_{k} \ \gamma_k \\ s.t. \quad & \left\|\mathbf{Z}^{-1}\mathbf{W}\right\|_F^2 \leq p_0 \end{aligned} \tag{5.14}$$

based on the fact that the power constraint is always active for $\mathscr{P}_{5.3}$. Accordingly, in the following we solve $\mathscr{P}_{5.3}$ instead. Compared to generic FD SINR balancing problem in [40, 49], the power constraint for $\mathscr{P}_{5.3}$ is on $\mathbf{Z}^{-1}\mathbf{W}$ instead of $\mathbf{W}$ only. Thanks to the fact that $\mathbf{Z}^{-1}$ is regarded as fixed when op-

timising $\mathbf{W}$, $\mathscr{P}_{5.3}$ can be solved by either the iterative algorithm in [49] or the bisection search approach in [40]. In this chapter, we employ the bisection search method to solve $\mathscr{P}_{5.3}$, based on the fact that SINR balancing and power minimisation problems are inverse problems [40]. To be more specific, we first express the power minimisation problem that corresponds to the SINR balancing problem $\mathscr{P}_{5.3}$ as a SOCP optimisation problem, given by

$$\begin{aligned} \mathscr{P}_{5.4}: \ & \min_{\mathbf{W}} \ \hat{p} \\ s.t. \quad & A_k(\mathbf{W}) \leq 0, \ \forall k \in \mathscr{K} \\ & C(\mathbf{W}) \leq 0 \end{aligned} \tag{5.15}$$

where $\mathscr{K} = \{1, 2, \cdots, K\}$ and each $A_k(\mathbf{W})$ is given by

$$A_k(\mathbf{W}) = \left\| \mathbf{W}^H \mathbf{H}^H \mathbf{I}_k \right\|_F + \sigma - \sqrt{1 + \frac{1}{\gamma_0}} \cdot [\mathbf{H}\mathbf{W}]_{k,k}, \ \forall k \in \mathscr{K}. \tag{5.16}$$

In (5.16), $\gamma_0$ is the SINR target for the corresponding power minimisation problem, and $\mathbf{I}_k$ denotes the $k$-th column in the identity matrix $\mathbf{I}$. Without loss of generality, we have assumed that $[\mathbf{D}\mathbf{W}]_{k,k} > 0$, as a phase rotation on each $\mathbf{w}_k$ will not have an impact on the resulting SINR expression. $C(\mathbf{W})$ is given by

$$C(\mathbf{W}) = \left\| \mathbf{Z}^{-1} \mathbf{W} \right\|_F - \hat{p}, \tag{5.17}$$

where $\hat{p}$ is the power variable to be optimised for the corresponding power minimisation problem. We further denote this power minimisation problem as $\mathscr{P}_{5.4}(\gamma_0)$. $\mathscr{P}_{5.4}$ is a typical SOCP problem and can be efficiently solved by power control algorithms or the semidefinite optimisation [39].

As the SINR balancing optimisation $\mathscr{P}_{5.3}$ and the power minimisation problem $\mathscr{P}_{5.4}(\gamma_0)$ are inverse problems, we can solve $\mathscr{P}_{5.3}$ by iteratively solving the corresponding power minimisation problems for different SINR target $\gamma_0$. Due to the inversion property, if $\mathscr{P}_{5.4}(\gamma_0) = p_0$, then its solution is also the optimal solution for $\mathscr{P}_{5.3}$ [40]. Based on the above description, we

**Algorithm 5.1** Bisection Search Algorithm for Solving $\mathscr{P}_{5.3}$

**input** : $\mathbf{H}$, $\mathbf{Z}^{-1}$, $\gamma_{max}$, $\gamma_{min}$, $p_0$
**output** : $\mathbf{W}$
**repeat**
$\gamma_0 \leftarrow \frac{\gamma_{\max}+\gamma_{\min}}{2}$;
$\hat{p} = \mathscr{P}_{5.4}(\gamma_0)$;
**if** $\hat{p}^2 \leq p_0$ **then**
    $\gamma_{\min} \leftarrow \gamma_0$;
**else**
    $\gamma_{\max} \leftarrow \gamma_0$;
**end if**
**until** $\hat{p}^2 = p_0$
Output $\mathbf{W}$.

present the following algorithm to efficiently solve $\mathscr{P}_{5.3}$, summarised in Algorithm 5.1, where $\gamma_{min}$ and $\gamma_{max}$ define the range of the target SINR.

### 5.3.2 Solving $\mathbf{z}_L$ When W is Fixed

When $\mathbf{W}$ is fixed, $\mathscr{P}_{5.1}$ is transformed into

$$\begin{aligned} \mathscr{P}_{5.5}: \ & \max_{\mathbf{z}_L} \min_{k} \ \gamma_k \\ s.t. \quad & \Re(z_{L_i}) \geq 0, \ \forall i \in \mathscr{N} \\ & f = \left\|\mathbf{Z}^{-1}\mathbf{W}\right\|_F \end{aligned} \tag{5.18}$$

Based on the expression of $\gamma_k$ in (5.11), it is observed that the noise amplification factor $f$ is the only variable when $\mathbf{W}$ is fixed. Therefore, $\mathscr{P}_{5.5}$ is equivalent to the following optimisation problem that minimises the noise amplification factor $f$, given by

$$\begin{aligned} \mathscr{P}_{5.6}: \ & \min_{\mathbf{z}_L} \ \left\|\mathbf{Z}^{-1}\mathbf{W}\right\|_F \\ s.t. \quad & \Re(z_{L_i}) \geq 0, \ \forall i \in \mathscr{N} \end{aligned} \tag{5.19}$$

To solve $\mathscr{P}_{5.6}$, we first study the inverse of the MC matrix $\mathbf{Z}^{-1}$. Based on (5.3), $\mathbf{Z}^{-1}$ is further obtained as

$$\begin{aligned}\mathbf{Z}^{-1} &= \left\{ [z_A \cdot \mathbf{I} + diag(\mathbf{z}_L)] [\mathbf{Z}_M + diag(\mathbf{z}_L)]^{-1} \right\}^{-1} \\ &= [\mathbf{Z}_M + diag(\mathbf{z}_L)] \cdot diag(\mathbf{z}_T),\end{aligned} \tag{5.20}$$

where

$$z_i = z_A + z_{L_i},\ \forall i \in \mathscr{N},\ \mathbf{z}_T = \left[ \frac{1}{z_1}, \frac{1}{z_2}, ..., \frac{1}{z_{N_t}} \right]^T. \tag{5.21}$$

By expanding (5.20), $\mathbf{Z}^{-1}$ can be expressed as

$$\mathbf{Z}^{-1} = \begin{bmatrix} 1 & \frac{z_{M_1}}{z_2} & \frac{z_{M_2}}{z_3} & \cdots & \frac{z_{M_{N_t-1}}}{z_{N_t}} \\ \frac{z_{M_1}}{z_1} & 1 & \frac{z_{M_1}}{z_3} & \ddots & \vdots \\ \frac{z_{M_2}}{z_1} & \frac{z_{M_1}}{z_2} & \ddots & \ddots & \frac{z_{M_2}}{z_{N_t}} \\ \vdots & \ddots & \ddots & \ddots & \frac{z_{M_1}}{z_{N_t}} \\ \frac{z_{M_{N_t-1}}}{z_1} & \cdots & \frac{z_{M_2}}{z_{N_t-2}} & \frac{z_{M_1}}{z_{N_t-1}} & 1 \end{bmatrix}. \tag{5.22}$$

By further introducing

$$\theta_i = \frac{z_{M_1}}{z_i},\ \forall i \in \mathscr{N}, \tag{5.23}$$

$\mathbf{Z}^{-1}$ in (5.22) can be decomposed into

$$\mathbf{Z}^{-1} = \mathbf{B}\mathbf{\Theta} + \mathbf{I}, \tag{5.24}$$

where $\mathbf{\Theta}$ is a diagonal matrix and $\mathbf{\Theta} = diag\left( [\theta_1, \theta_2, \cdots, \theta_{N_t}]^T \right)$. The matrix $\mathbf{B}$ is given by

$$\mathbf{B} = \begin{bmatrix} 0 & 1 & \frac{z_{M_2}}{z_{M_1}} & \cdots & \frac{z_{M_{N_t-1}}}{z_{M_1}} \\ 1 & 0 & 1 & \ddots & \vdots \\ \frac{z_{M_2}}{z_{M_1}} & 1 & \ddots & \ddots & \frac{z_{M_2}}{z_{M_1}} \\ \vdots & \ddots & \ddots & \ddots & 1 \\ \frac{z_{M_{N_t-1}}}{z_{M_1}} & \cdots & \frac{z_{M_2}}{z_{M_1}} & 1 & 0 \end{bmatrix}. \tag{5.25}$$

Based on (5.21) and (5.23), we express the load impedance $z_{L_i}$ as

$$z_{L_i} = \frac{z_{m_1}}{\theta_i} - z_A,\ \forall i \in \mathcal{N}, \tag{5.26}$$

and the constraint that $\Re(z_{L_i}) \geq 0$ can be further transformed into

$$\Re\left(\frac{z_{m_1}}{\theta_i}\right) \geq \Re(z_A) \ \Rightarrow\ \Re(\theta_i)\Re(z_{m_1}) + \Im(\theta_i)\Im(z_{m_1}) \geq |\theta_i|^2 \cdot \Re(z_A). \tag{5.27}$$

(5.27) is convex, where we note that a practical antenna has $\Re(z_A) > 0$ [63]. By substituting (5.24) and (5.27) into $\mathscr{P}_{5.6}$, the optimisation problem is finally formulated as

$$\begin{aligned} \mathscr{P}_{5.7}:\ & \min_{\boldsymbol{\Theta}} \ \|\mathbf{B}\boldsymbol{\Theta}\mathbf{W} + \mathbf{W}\|^2 \\ s.t.\quad & \Re(\theta_i)\Re(z_{m_1}) + \Im(\theta_i)\Im(z_{m_1}) \geq |\theta_i|^2 \cdot \Re(z_A),\ \forall i \in \mathcal{N} \end{aligned} \tag{5.28}$$

which is a least-squares problem and can be efficiently solved with convex optimisation tools. Subsequently, the optimal load value $z_{L_i}$ can be obtained by (5.26), and the resulting optimal mutual coupling matrix is obtained as

$$\mathbf{Z}^* = (\mathbf{B}\boldsymbol{\Theta}^* + \mathbf{I})^{-1}. \tag{5.29}$$

### 5.3.3 Joint Iterative Algorithm

Based on the above description, the proposed joint AD precoding method can be obtained by alternately optimising $\mathbf{W}$ and $\mathbf{z}_L$ until convergence or a maximum number of iterations is reached. We summarise the alternating optimisation algorithm for the joint scheme in Algorithm 5.2, where $\mathbf{z}_L^0$ is the initial load impedance vector, and $N_{max}$ denotes the maximum iteration number. $\delta$ is a variable that represents the convergence accuracy, which is defined as $\delta = \left\|\mathbf{W}^{(n+1)} - \mathbf{W}^{(n)}\right\|_F$, or equivalently $\delta = \left\|\mathbf{Z}^{(n+1)} - \mathbf{Z}^{(n)}\right\|_F$. $\delta_{th}$ defines the accuracy threshold.

### 5.3.4 Convergence Discussion

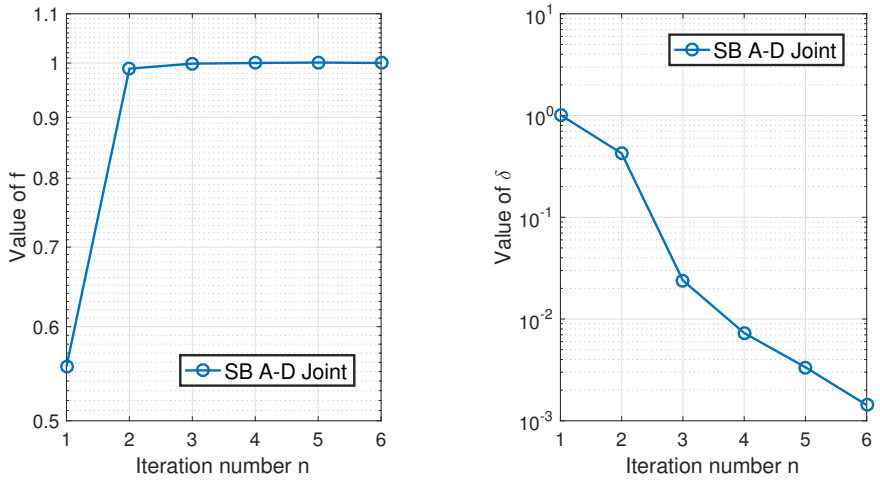
It is observed in Algorithm 5.2 that the two sub-problems have the same objective function $\mathscr{P}_{5.1}$ within each iteration, which is then transformed into equivalent sub-problems $\mathscr{P}_{5.3}$ and $\mathscr{P}_{5.5}$. Moreover, the sub-problem $\mathscr{P}_{5.3}$ to obtain $\mathbf{W}^{(n)}$ and the sub-problem $\mathscr{P}_{5.5}$ to obtain $\mathbf{Z}^{(n)}$ are optimally solved by Algorithm 5.1 and $\mathscr{P}_{5.7}$ for that iteration, respectively. Therefore, Algorithm 5.2 is guaranteed to converge [69, 121], and it is easy to observe that when the global optimum is reached, $f^* = \sqrt{M}$. More importantly, it is found that the convergence speed of the proposed joint algorithm is fast, which is shown below in Fig. 5.2, where we depict the value of both $f$ and $\delta$ with respect to the increasing iteration number $n$, and for simplicity we assume that $M = 1$. 'SB A-D Joint' denotes the proposed joint AD SINR balancing optimisation. In Fig. 5.2, the initial antenna load vector $\mathbf{z}_L^0$ is $z_{L_i} = 50\Omega$, $\forall i \in \mathscr{N}$, $\gamma_{min} = 0$, and $\gamma_{max}$ is selected as $\gamma_{\max} = \frac{\max\left(\|\mathbf{h}_k\|^2\right)}{\sigma^2}$ [40]. It can be observed that the proposed algorithm is convergent within $n = 4$ iterations.

**Algorithm 5.2** Joint Iterative Algorithm for Solving $\mathscr{P}_{5.1}$

**input** : $\mathbf{H}$, $\mathbf{z}_L^0$, $N_{\max}$, $\delta_{th}$
**output** : $\mathbf{P}^*$
$n = 0$;
$\mathbf{Z}^0 = \mathbf{Z}\left(\mathbf{z}_L^0\right)$, $\mathbf{W}^0 = \mathbf{0}$;
**while** $n \leq N_{max}$ and $\delta \geq \delta_{th}$ **do**
  Obtain $\mathbf{W}^{(n+1)}$ by Algorithm 5.1 with $\mathbf{Z}^{(n)}$;
  Obtain $\mathbf{Z}^{(n+1)}$ by solving $\mathscr{P}_{5.7}$ with $\mathbf{W}^{(n+1)}$;
  $\delta = \left\|\mathbf{W}^{(n+1)} - \mathbf{W}^{(n)}\right\|$;
  $n = n + 1$;
**end while**
$\mathbf{W}^* = \mathbf{W}^{(n)}$, $\mathbf{Z}^* = \left(\mathbf{B}\boldsymbol{\Theta}^{(n)} + \mathbf{I}\right)^{-1}$;
Calculate $f^*$;
Output $\mathbf{P}^* = \frac{1}{f^*} \cdot (\mathbf{Z}^*)^{-1}\mathbf{W}^*$.

## 5.4 Low-Complexity Decoupled Approach

Due to the fact that $\mathbf{Z}^{-1}$ and $\mathbf{W}$ need to be updated alternately, the proposed joint algorithm above involves considerable computational costs. In




**(a)** Value of $f$ **(b)** Value of $\delta$

**Figure 5.2:** Convergence speed of the joint iterative approach, $N_t = K = 4$, $N_{\max} = 6$, $\delta_{th} = 10^{-3}$

this chapter, we further explore a low-complexity alternative to the joint optimisation above, where a more practical decoupled method is introduced. In practice, closed-form ZF and RZF precoding approaches are widely employed due to their complexity benefits. While the joint AD precoding method in Chapter 5.3 does not directly apply to these closed-form approaches, the decoupled method introduced in this chapter enables the combination of the MC exploitation with such closed-form precoders, which leads to a much reduced computational cost.

For the proposed decoupled technique, we still construct $\mathbf{P}$ as in (5.8), and we denote $\mathbf{W}_d$ as the digital precoding matrix for the decoupled approach. In the decoupled design, $\mathbf{W}_d$ is firstly obtained with conventional FD precoding approaches assuming no coupling effect, and subsequently we design $\mathbf{z}_L$ to minimise the noise amplification factor $f$. With the proposed decoupled approach, the computational complexity is greatly reduced, while it will be observed that there is only a slight performance loss for SINR balancing optimisation, compared to the joint approach. The decoupled precoding method is summarised in Algorithm 5.3 below.

**Algorithm 5.3** Decoupled AD Precoding Approach

**input** : $\mathbf{H}$
**output** : $\mathbf{P}_d^*$
Obtain $\begin{cases} \mathbf{W}_d^* \text{ by } \mathscr{P}_{5.2} \text{ with } \mathbf{Z}=\mathbf{I}, \text{ for SINR balancing} \\ \mathbf{W}_d^* = \mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H\right)^{-1}, \text{ for closed}-\text{form ZF} \end{cases}$;
Obtain $\mathbf{Z}^*$ by solving $\mathscr{P}_{5.7}$ with the obtained $\mathbf{W}_d^*$;
Calculate $f_d^*$;
Output $\mathbf{P}_d^* = \frac{1}{f_d^*} \cdot \left(\mathbf{Z}_d^*\right)^{-1}\mathbf{W}_d^*$.

## 5.5 Sequential Search Method for Quantised Loads

In the above sections, while we have considered the realistic constraint for the varactors where $\Re\left(z_{L_i}\right) \geq 0, \forall i \in \mathscr{N}$, this constraint is from the perspective of enabling the radiation of the antenna array. In optimisation problems $\mathscr{P}_{5.1}$ - $\mathscr{P}_{5.7}$, we have assumed an infinite precision for the varactors, i.e., each $z_{L_i}$ can have any arbitrary continuous values. This assumption may not be feasible in practical deployments, where hardware components with finite precision are used. Therefore, we proceed to consider a more practical case where only quantised versions of the analog loads are available, and present a practical sequential search method based on the greedy algorithm for the analog precoding optimisation. In this case, the constraint for the tunable load is modified into

$$z_{L_i} \in \mathscr{U}, \ \forall i \in \mathscr{N}, \tag{5.30}$$

where $\mathscr{U}$ denotes the set that consists of quantised load values with a finite number. For the precoding method with limited-precision antenna loads, the optimisation problem $\mathscr{P}_{5.6}$ in (5.19) for $\mathbf{z}_L$ can be reformulated as

$$\begin{aligned} \mathscr{P}_{5.8}: \ & \min_{\mathbf{z}_L} \ \left\|\mathbf{Z}^{-1}\mathbf{W}\right\|_F \\ s.t. \quad & z_{L_i} \in \mathscr{U}, \ \forall i \in \mathscr{N} \end{aligned} \tag{5.31}$$

We denote each potential load value in $\mathscr{U}$ as $z_m^Q$, and we further denote $D_R$ and $D_I$ as the total number of potential values for the real and imaginary part of each $z_m^Q \in \mathscr{U}$, respectively. Accordingly, the cardinality of $\mathscr{U}$ can be obtained as

$$card\left(\mathscr{U}\right) = D_R D_I = D, \tag{5.32}$$

based on which we express the real and imaginary part of each $z_m^Q \in \mathscr{U}$ as

$$\Re\left(z_m^Q\right) \in \left\{z_1^{\Re}, z_2^{\Re}, \cdots, z_{D_R}^{\Re}\right\},\ \Im\left(z_m^Q\right) \in \left\{z_1^{\Im}, z_2^{\Im}, ..., z_{D_I}^{\Im}\right\},\ \forall m \in \mathscr{N}, \tag{5.33}$$

where the quantised potential load values $z_n^{\Re}$ and $z_n^{\Im}$ in (5.33) can be obtained as

$$\begin{aligned} z_n^{\Re} &= \frac{n \cdot R_{upper}}{D_R},\ \forall n \in \{1, 2, \cdots, D_R\}; \\ z_n^{\Im} &= I_{lower} + \frac{n \cdot (I_{upper} - I_{lower})}{D_I},\ \forall n \in \{1, 2, \cdots, D_I\}. \end{aligned} \tag{5.34}$$

In (5.34), $R_{upper}$ denotes the upper bound for the real part of each $z_m^Q \in \mathscr{U}$ where the constraint that $\Re\left(z_{L_i}\right) \geq 0, \forall i \in \mathscr{N}$ has already been included implicitly, while $I_{lower}$ and $I_{upper}$ denote the lower bound and upper bound for the imaginary part of $z_m^Q \in \mathscr{U}$. We note that here we also constrain the imaginary part of each $z_{L_i}$ in a finite range for practical consideration.

It is observed that the optimal quantised load vector can only be obtained via an exhaustive search over a total number of $D$ possible load values for each antenna element, which is highly computationally expensive, especially when the number of transmit antennas $N_t$ is large or the precision of the tunable loads is high. While it is also feasible to quantise the optimal load values obtained by the optimisation, this approach still requires solving an optimisation problem, which may also be computationally inefficient in practice. Accordingly, to obtain a feasible solution efficiently, in this chapter we propose a low-complexity sequential search approach based on a greedy algorithm. To be specific, for the load value of each antenna element, we sequentially select the load value that provides the highest performance

improvements, which is achieved by selecting the load value that generates the lowest value of $f$, while assuming the loads on other antenna elements fixed. The proposed scheme is summarised in Algorithm 5.4, which can be employed upon the joint precoding approach in Chapter 5.3 or the decoupled method in Chapter 5.4 to efficiently obtain a feasible quantised load value for each antenna element.

**Remark on the Performance of the Sequential Search Approach:** While we do not guarantee that the proposed sequential search scheme can achieve the optimal performance, our proposed approach indeed offers a low-complexity solution that achieves performance gains over fixed coupling effect for the practical scenarios, where quantised tunable loads are employed and the exhaustive search is inapplicable due to the high complexity. Moreover, it will be further shown in the numerical results that for a moderate-size MIMO system with a small quantisation level $D$, the proposed approach indeed achieves a close-to-optimal performance.

Furthermore, to numerically show the complexity benefits of the proposed algorithm, we evaluate the computational costs in terms of the total number of possible load combinations to visit for both the exhaustive search and the proposed sequential search, obtained as

$$\begin{aligned} N_{exhaustive} &= D^{N_t} = (D_R D_I)^{N_t}, \\ N_{sequential} &= DN_t = D_R D_I N_t. \end{aligned} \tag{5.35}$$

An illustrative example is presented in Table 5.1 with some typical values of the antenna number and quantisation level, where we can observe a significant complexity gain for the proposed sequential search method, compared to the exhaustive search scheme which requires a high computational cost.

**Algorithm 5.4** Sequential Search Scheme for Quantised Loads based on the Greedy Algorithm

**input** : $\mathbf{z}_L^0$, $\mathscr{U}$, $\mathbf{W}$
**output** : $\mathbf{Z}^*$
Obtain $\mathbf{Z}_0$ with $\mathbf{z}_L^0$ by (5.3), and $f_0 = \left\|\mathbf{Z}_0^{-1}\mathbf{W}\right\|_F$;
**for** $m = 1 : N_t$ **do**
  $z_t = \mathbf{z}_L^0(m)$, $\mathbf{z}_L = \mathbf{z}_L^0$;
  **for** $n = 1 : D$ **do**
    $\mathbf{z}_L(m) \leftarrow z_n^Q$;
    Obtain $\mathbf{Z}$ with $\mathbf{z}_L$ by (5.3), and $f = \left\|\mathbf{Z}^{-1}\mathbf{W}\right\|_F$;
    **if** $f < f_0$ **then**
      $f_0 \leftarrow f$, $z_t \leftarrow z_n^Q$;
    **end if**
  **end for**
  $\mathbf{z}_L(m) \leftarrow z_t$;
**end for**
Obtain $\mathbf{Z}^*$ with $\mathbf{z}_L$ by (5.3).

| Exhaustive Search Scheme | | | | | |
|---|---|---|---|---|---|
| Antenna Number | Quantisation Level | | | | |
| | 2 | 4 | 6 | 8 | 10 |
| 4 | 256 | $6.6 \times 10^4$ | $1.7 \times 10^6$ | $1.7 \times 10$ | $1 \times 10^8$ |
| 8 | $6.6 \times 10^4$ | $4.3 \times 10^9$ | $2.8 \times 10^{12}$ | $2.8 \times 10^{14}$ | $1 \times 10^{16}$ |
| 16 | $4.3 \times 10^9$ | $1.8 \times 10^{19}$ | $8 \times 10^{24}$ | $7.9 \times 10^{28}$ | $1 \times 10^{32}$ |
| Sequential Search Scheme | | | | | |
| Antenna Number | Quantisation Level | | | | |
| | 2 | 4 | 6 | 8 | 10 |
| 4 | 16 | 64 | 144 | 256 | 400 |
| 8 | 32 | 128 | 288 | 512 | 800 |
| 16 | 64 | 256 | 576 | 1024 | 1600 |

**Table 5.1:** Number of load combinations to visit for the exhaustive search and proposed sequential search scheme

## 5.6 Performance Analysis and Optimality for Closed-Form Precoding Methods

In this chapter, we conduct performance analyses in terms of the received SINR for imperfect CSI, followed by the optimality demonstration of the proposed MC exploitation for closed-form precoders in imperfect CSI.

### 5.6.1 Estimation of H

We observe that the BS needs to obtain $\mathbf{H}$ to employ the proposed approaches. At the BS, the CSI is obtained with channel estimation techniques based on the duplex mode, which has been introduced in Chapter 2.3.4. Regardless of which duplex mode is employed, we can only obtain $\mathbf{D} = \mathbf{HZ}$ with standard channel estimation methods. Therefore, first we need to extract $\mathbf{H}$ from $\mathbf{HZ}$ to facilitate the proposed approaches. A similar approach to Chapter 4.5.1 can be employed to extract $\mathbf{H}$ from the channel matrix $\mathbf{D}$ by estimating the channel with a reference coupling matrix $\mathbf{Z}_0$, and we omit the details in this chapter for brevity. With $\mathbf{Z}_0$ known to the BS, $\mathbf{H}$ can be extracted from $\mathbf{HZ}_0$ and expressed as

$$\mathbf{H} = \mathbf{D}\mathbf{Z}_0^{-1}. \tag{5.36}$$

With $\mathbf{H}$ obtained, the proposed approaches can be employed during data transmission.

### 5.6.2 Perfect CSI

In the case of perfect CSI, the received SINR of user $k$ for SINR balancing can be obtained by (5.11), based on which we obtain the sum rate as

$$R = \sum_{k=1}^{K} \log_2 (1 + \gamma_k). \tag{5.37}$$

We proceed to calculate the received SINR for imperfect CSI, where we employ the statistical CSI error model as an example, and the extension to the norm-bounded CSI can be similarly obtained.

### 5.6.3 Imperfect CSI - Statistical CSI Errors

Based on (5.36) and the imperfect CSI model for TDD systems in (2.53), which is described in Chapter 2.3.4, we can further express $\mathbf{H}$ as

$$\begin{aligned}\mathbf{H} &= \alpha \cdot \left(\hat{\mathbf{H}} + \mathbf{E}\mathbf{Z}^{-1}\right) + \mathbf{Q}\mathbf{Z}^{-1} \\ &= \alpha \cdot \hat{\mathbf{H}} + (\alpha \cdot \mathbf{E} + \mathbf{Q})\,\mathbf{Z}^{-1}.\end{aligned} \tag{5.38}$$

We proceed to study the received SINR for imperfect CSI. Based on (5.1) and (5.38), the received signal vector $\hat{\mathbf{y}}$ with the proposed approach can be expressed as

$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{D}\hat{\mathbf{x}} + \mathbf{n} \\ &= \left[\alpha \cdot \hat{\mathbf{H}} + (\alpha \cdot \mathbf{E} + \mathbf{Q})\mathbf{Z}^{-1}\right]\mathbf{Z} \cdot \frac{1}{\hat{f}} \cdot \mathbf{Z}^{-1}\hat{\mathbf{W}}\mathbf{s} + \mathbf{n} \\ &= \frac{\alpha}{\hat{f}} \cdot \hat{\mathbf{H}}\hat{\mathbf{W}}\mathbf{s} + \frac{1}{\hat{f}} \cdot (\alpha \cdot \mathbf{E} + \mathbf{Q})\,\mathbf{Z}^{-1}\hat{\mathbf{W}}\mathbf{s} + \mathbf{n},\end{aligned} \tag{5.39}$$

In the presence of statistical CSI errors, each receiver needs to scale the received signal by $\frac{\hat{f}}{\alpha}$, and the resulting received symbol vector can be expressed as

$$\hat{\mathbf{r}} = \hat{\mathbf{H}}\hat{\mathbf{W}}\mathbf{s} + \left(\mathbf{E} + \frac{1}{\alpha} \cdot \mathbf{Q}\right)\mathbf{Z}^{-1}\hat{\mathbf{W}}\mathbf{s} + \frac{\hat{f}}{\alpha} \cdot \mathbf{n}. \tag{5.40}$$

The second term in (5.40) introduces an additional noise term due to the CSI estimation errors. We define the equivalent noise term as

$$\hat{\mathbf{n}} = \left(\mathbf{E} + \frac{1}{\alpha} \cdot \mathbf{Q}\right)\mathbf{Z}^{-1}\hat{\mathbf{W}}\mathbf{s} + \frac{\hat{f}}{\alpha} \cdot \mathbf{n}. \tag{5.41}$$

Based on the derivation in [34], it is shown that the distribution of $\hat{\mathbf{n}}$ conditioned on $\mathbf{Z}$ and $\mathbf{s}$ is i.i.d. Gaussian with zero mean, i.e. $\hat{\mathbf{n}} \sim \mathbb{CN}(\mathbf{0}, \upsilon \cdot \mathbf{I})$, with the equivalent noise power given by

$$\upsilon = \left[\eta + \frac{(1-\alpha^2)}{\alpha^2}\right] \cdot \hat{f}^2 + \frac{\hat{f}^2\sigma^2}{\alpha^2} = \frac{\hat{f}^2}{\alpha^2}\left[\alpha^2\eta + (1-\alpha^2) + \sigma^2\right], \tag{5.42}$$

where we note that $\hat{\mathbf{x}}$ is normalised, which leads to $\|\hat{\mathbf{x}}\|_F^2 = 1$ and $\left\|\mathbf{Z}^{-1}\hat{\mathbf{W}}\mathbf{s}\right\|_F^2 = \hat{f}^2$. It is further obtained that the derivation in (5.39)-(5.42) is independent of the precoding approaches applied in the digital domain. Accordingly, we can express the resulting received SINR for the $k$-th user in the presence of statistical CSI errors as

$$\hat{\gamma}_k = \frac{\left|\hat{\mathbf{h}}_k \hat{\mathbf{w}}_k^*\right|^2}{\sum\limits_{i\neq k}^{K} \left|\hat{\mathbf{h}}_k \hat{\mathbf{w}}_i^*\right|^2 + \frac{\hat{f}^2}{\alpha^2}\left[\alpha^2\eta + (1-\alpha^2) + \sigma^2\right]}, \tag{5.43}$$

and the sum rate can be calculated by (5.37) via the obtained $\hat{\gamma}_k$. Based on the observation in (5.43), we note the optimality of the proposed AD techniques for closed-form digital precoding approaches with imperfect CSI.

**Optimality for Closed-Form Precoding Methods:** When closed-form precoding methods are employed in the digital domain, each $\hat{\mathbf{w}}_k^*$, $\forall k \in \mathscr{K}$ in (5.43) is only dependent on the CSI and is fixed. The received SINR of the $k$-th user with statistical CSI errors is therefore only relevant to the noise amplification factor $\hat{f}$. For the proposed analog precoding approaches that minimise the noise amplification factor $\hat{f}$, the power of the equivalent noise vector $\hat{\mathbf{n}}$ for the imperfect CSI scenarios is also minimised, which means that the solutions obtained via $\mathscr{P}_{5.1}$ - $\mathscr{P}_{5.8}$ assuming perfect CSI are also optimal for the corresponding cases in the presence of statistical CSI errors, and the proposed methods can be directly extended to such cases.

### 5.6.4 Imperfect CSI - Norm-Bounded CSI Errors

The norm-bounded CSI error model for FDD systems is given in (2.56), as described in Chapter 2.3.4. The **optimality for closed-form precoding approaches** trivially extends to the case of norm-bounded CSI errors when closed-form precoding approaches are employed in the digital domain. In such case, instead of obtaining the exact analytical SINR expression, we can derive the lower bound of the received SINR based on a similar derivation

in (5.39)-(5.43), given by

$$\hat{\gamma}_k \geq \frac{\left|\hat{\mathbf{h}}_k \hat{\mathbf{w}}_k^*\right|^2}{\sum\limits_{i \neq k}^{K} \left|\hat{\mathbf{h}}_k \hat{\mathbf{w}}_i^*\right|^2 + \hat{f}^2 \sigma^2 + \delta_k} = \hat{\gamma}_k^{lower}, \ \forall k \in \mathscr{K}. \tag{5.44}$$

## 5.7 Proposed Robust Precoding for SINR Optimisation

In this chapter, the robust designs for SINR optimisation for both the statistical CSI error and the norm-bounded CSI error model are proposed.

### 5.7.1 Robust Design - Statistical CSI Errors

Based on the derivation in (5.39)-(5.43), the received SINR with statistical CSI errors can be regarded similar to the perfect CSI case with an equivalent noise term $\hat{\mathbf{n}}$. It is further noted that the power of $\hat{\mathbf{n}}$ is independent of the coupling matrix $\mathbf{Z}$. Therefore, for the robust SINR balancing optimisation with statistical CSI errors, instead of employing the noise power of $\mathbf{n}$, we apply the power of the equivalent noise $\hat{\mathbf{n}}$ for $A_k(\mathbf{W})$ in (5.15), and we express $\hat{A}_k(\mathbf{W})$ for the robust precoding as

$$\hat{A}_k(\mathbf{W}) = \left\|\mathbf{W}^H \mathbf{H}^H \mathbf{I}_k\right\|_F + \hat{\sigma} - \sqrt{1 + \frac{1}{\gamma_0}} \cdot [\mathbf{HW}]_{k,k}, \ \forall k \in \mathscr{K}, \tag{5.45}$$

where $\hat{\sigma}$ denotes the power of the equivalent noise $\hat{\mathbf{n}}$ before rescaling and can be obtained based on (5.42) as

$$\hat{\sigma} = \sqrt{\frac{1}{\alpha^2}\left[\alpha^2 \eta + (1-\alpha^2) + \sigma^2\right]}. \tag{5.46}$$

By substituting $\hat{A}_k(\mathbf{W})$ obtained in (5.45) into $\mathscr{P}_{5.4}$, the robust SINR balancing optimisation with statistical CSI errors can be solved.

### 5.7.2 Robust Design - Norm-Bounded CSI Errors

We further consider the robust SINR optimisation with norm-bounded CSI errors, which is formulated into the following optimisation problem

$$\begin{aligned} \mathscr{P}_{5.9}: \ & \max_{\hat{\mathbf{z}}_L, \hat{\mathbf{W}}} \min_{k} \ \hat{\gamma}_k \\ s.t. \quad & \Re\left(\hat{z}_{L_i}\right) \geq 0, \ \forall i \in \mathscr{N} \\ & \hat{\gamma}_k = \frac{\left|\mathbf{h}_k \hat{\mathbf{w}}_k\right|^2}{\sum\limits_{i \neq k}^{K} \left|\mathbf{h}_k \hat{\mathbf{w}}_i\right|^2 + \hat{f}^2 \sigma^2}, \ \mathbf{h}_k \in \mathscr{H}_k, \ , \forall k \in \mathscr{K} \end{aligned} \tag{5.47}$$

Similar to the case of perfect CSI, we consider the alternating optimisation approach, and in the following we first assume a fixed mutual coupling matrix and solve the robust SINR optimisation for imperfect CSI by solving a sequence of equivalent power minimisation problems with the bisection search approach [122, 123]. Accordingly, in the following we focus on solving the corresponding robust power minimisation problem for norm-bounded CSI errors, and the optimal precoding matrix for the robust SINR balancing can be obtained via the bisection search in Algorithm 5.1. For simplicity and without loss of generality we assume $p_0 = 1$, and the corresponding power minimisation problem for $\mathscr{P}_{5.9}$ can be formulated into

$$\begin{aligned} \mathscr{P}_{5.10}: \ & \min_{\hat{\mathbf{W}}} \ \hat{p} \\ s.t. \quad & \hat{p} \geq \left\| \hat{\mathbf{Z}}^{-1} \hat{\mathbf{W}} \right\|_F^2 \\ & \frac{\left|\mathbf{h}_k \hat{\mathbf{w}}_k\right|^2}{\sum\limits_{i \neq k}^{K} \left|\mathbf{h}_k \hat{\mathbf{w}}_i\right|^2 + \sigma^2} \geq \hat{\gamma}_0, \ \mathbf{h}_k \in \mathscr{H}_k, \ \forall k \in \mathscr{K} \end{aligned} \tag{5.48}$$

Note that the optimisation problem $\mathscr{P}_{5.10}$ contains infinite number of constraints and is difficult to be directly solved. Therefore, we propose to consider the worst-case received SINR for user $k$ within the uncertainty region

$\mathscr{H}_k$, defined as

$$\widetilde{SINR}_k = \min_{\mathbf{h}_k \in \mathscr{H}_k} \frac{|\mathbf{h}_k \hat{\mathbf{w}}_k|^2}{\sum\limits_{i \neq k}^{K} |\mathbf{h}_k \hat{\mathbf{w}}_i|^2 + \sigma^2}, \tag{5.49}$$

and the optimisation problem $\mathscr{P}_{5.10}$ is equivalent to

$$\begin{aligned} \mathscr{P}_{5.11}: \ & \min_{\hat{\mathbf{W}}} \ \hat{p} \\ s.t. \quad & \hat{p} \geq \left\| \hat{\mathbf{Z}}^{-1} \hat{\mathbf{W}} \right\|_F^2 \\ & \widetilde{SINR}_k \geq \hat{\gamma}_0, \ \mathbf{h}_k \in \mathscr{H}_k, \ \forall k \in \mathscr{K} \end{aligned} \tag{5.50}$$

With the definition of $\widetilde{SINR}_k$ and $\mathscr{H}_k$, we further denote

$$\mathbf{T}_k = \frac{1}{\hat{\gamma}_0} \hat{\mathbf{w}}_k \hat{\mathbf{w}}_k^H - \sum_{i \neq k} \hat{\mathbf{w}}_i \hat{\mathbf{w}}_i^H, \tag{5.51}$$

and the constraint that $\widetilde{SINR}_k \geq \hat{\gamma}_0$ can be further transformed into

$$\begin{aligned} & \mathbf{h}_k \mathbf{T}_k \mathbf{h}_k^H - \sigma^2 \geq 0 \\ \Rightarrow & \left(\hat{\mathbf{h}}_k + \mathbf{e}_k\right) \mathbf{T}_k \left(\hat{\mathbf{h}}_k + \mathbf{e}_k\right)^H - \sigma^2 \geq 0 \\ \Rightarrow & \mathbf{e}_k \mathbf{T}_k \mathbf{e}_k^H + \mathbf{e}_k \left(\mathbf{T}_k \hat{\mathbf{h}}_k^H\right) + \left(\mathbf{T}_k \hat{\mathbf{h}}_k^H\right)^H \mathbf{e}_k^H + \hat{\mathbf{h}}_k \mathbf{T}_k \hat{\mathbf{h}}_k^H - \sigma^2 \geq 0, \ \forall \mathbf{e}_k \mathbf{e}_k^H \leq \delta_k^2 \end{aligned} \tag{5.52}$$

**Lemma: S-procedure [124]:** Let $\mathbf{T} \in \mathbb{C}^{N_t \times N_t}$ be a complex Hermitian matrix, $\mathbf{b} \in \mathbb{C}^{N_t \times 1}$ and a scalar $c$. For a vector $\mathbf{v} \in \mathbb{C}^{1 \times N_t}$, the following condition

$$\mathbf{v}\mathbf{T}\mathbf{v}^H + \mathbf{v}\mathbf{b} + \mathbf{b}^H \mathbf{v}^H + c \geq 0, \ \forall \|\mathbf{v}\|^2 \leq r^2 \tag{5.53}$$

is true if and only if there exists a non-negative value $t$, such that the matrix $\mathbf{U}$ given below

$$\mathbf{U} = \begin{bmatrix} \mathbf{T} + t \cdot \mathbf{I} & \mathbf{b} \\ \mathbf{b}^H & c - tr^2 \end{bmatrix} \tag{5.54}$$

is positive semi-definite. By employing the S-procedure, (5.52) can be transformed into a positive semi-definite form, expressed as

$$\begin{bmatrix} \mathbf{T}_k + t_k \cdot \mathbf{I} & \mathbf{T}_k \hat{\mathbf{h}}_k^H \\ \hat{\mathbf{h}}_k \mathbf{T}_k^H & \hat{\mathbf{h}}_k \mathbf{T}_k \hat{\mathbf{h}}_k^H - \sigma^2 - t_k \delta_k \end{bmatrix} \succeq 0, \ \forall k \in \mathscr{K} \tag{5.55}$$

where $t_k \geq 0$ is an introduced auxiliary variable. By defining $\mathbf{X}_k = \hat{\mathbf{w}}_k \hat{\mathbf{w}}_k^H$, $\forall k \in \mathscr{K}$, $\left\| \hat{\mathbf{Z}}^{-1} \hat{\mathbf{W}} \right\|_F^2$ can be transformed into

$$\begin{aligned} \left\| \hat{\mathbf{Z}}^{-1} \mathbf{W} \right\|_F^2 &= tr\left\{ \mathbf{Z}^{-1} \hat{\mathbf{W}} \hat{\mathbf{W}}^H \left( \mathbf{Z}^{-1} \right)^H \right\} \\ &= tr\left\{ \mathbf{Z}^{-1} \left[ \hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, ..., \hat{\mathbf{w}}_K \right] \left[ \hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2, ..., \hat{\mathbf{w}}_K \right]^H \left( \mathbf{Z}^{-1} \right)^H \right\} \\ &= \sum_{k=1}^{K} tr\left\{ \mathbf{Z}^{-1} \mathbf{X}_k \left( \mathbf{Z}^{-1} \right)^H \right\}. \end{aligned} \tag{5.56}$$

Then, with (5.55) and (5.56), $\mathscr{P}_{5.11}$ can be recast as

$$\begin{aligned} \mathscr{P}_{5.12}: \ & \min_{\mathbf{X}_k, t_k} \ \hat{p} \\ s.t. \quad & \hat{p} \geq \sum_{k=1}^{K} tr\left\{ \mathbf{Z}^{-1} \mathbf{X}_k \left( \mathbf{Z}^{-1} \right)^H \right\} \\ & \mathbf{T}_k = \frac{1}{\gamma_k} \mathbf{X}_k - \sum_{i \neq k} \mathbf{X}_i, \ \forall k \in \mathscr{K} \\ & \begin{bmatrix} \mathbf{T}_k + t_k \cdot \mathbf{I} & \mathbf{T}_k \hat{\mathbf{h}}_k^H \\ \hat{\mathbf{h}}_k \mathbf{T}_k^H & \hat{\mathbf{h}}_k \mathbf{T}_k \hat{\mathbf{h}}_k^H - \sigma^2 - t_k \delta \end{bmatrix} \succeq 0, \ \forall k \in \mathscr{K} \\ & t_k \geq 0, \ \mathbf{X}_k \succeq 0, \ \text{rank}\left( \mathbf{X}_k \right) = 1, \ \forall k \in \mathscr{K} \end{aligned} \tag{5.57}$$

By dropping the rank constraint $\text{rank}\left( \mathbf{X}_k \right) = 1$, $\mathscr{P}_{5.12}$ becomes a standard semidefinite programming (SDP) optimisation and can be efficiently solved with existing optimisation tools. Furthermore, it has been shown in [125] that the optimal solution obtained satisfies $\text{rank}\left( \mathbf{X}_k^* \right) = 1$, when the uncertainty bound $\delta_k$ is small. In this case, the exact optimal solution for the original problem $\mathscr{P}_{5.12}$ can be obtained by employing the eigenvalue decomposition. On the other hand, if $\text{rank}\left( \mathbf{X}_k^* \right) > 1$, randomisation techniques in [126]

or rank reduction approaches in [127] that are widely adopted can be employed to obtain a feasible rank-one solution.

Specific to our optimisation problem, after the digital precoder is obtained via the bisection search scheme with Algorithm 5.1, the optimisation on $\hat{\mathbf{z}}_L$ can be employed by incorporating $\mathscr{P}_{5.12}$ into the joint and decoupled precoding approach, respectively. For the joint scheme, $\hat{\mathbf{W}}^{(n+1)}$ within each iteration is obtained by solving $\mathscr{P}_{5.12}$ with a fixed MC matrix $\mathbf{Z}^{(n)}$, where we note that an iterative design in Algorithm 5.2 is still needed. As for the decoupled scheme, the digital precoders are first obtained by solving $\mathscr{P}_{5.12}$ with $\mathbf{Z} = \mathbf{I}$, and the optimal load values that minimise the total transmit power are obtained by solving $\mathscr{P}_{5.7}$. In the case of quantised analog load values, the sequential search algorithm in Chapter 5.5 can be directly applied to the robust precoding methods. With the above approaches, the robust SINR optimisation with MC exploitation for norm-bounded CSI errors can be solved.

For the case where a fixed coupling matrix is considered with norm-bounded CSI errors, we obtain

$$\begin{aligned} \mathbf{d}_k &= \left(\hat{\mathbf{h}}_k + \mathbf{e}_k\right)\mathbf{Z}_0 = \hat{\mathbf{d}}_k + \mathbf{e}_k\mathbf{Z}_0 \\ &= \hat{\mathbf{d}}_k + \hat{\mathbf{e}}_k, \end{aligned} \tag{5.58}$$

where $\hat{\mathbf{e}}_k$ denotes the channel error vector for the fixed MC effect. Based on the sub-multiplicativity property of the norm, we can further obtain

$$\|\hat{\mathbf{e}}_k\|_F = \|\mathbf{e}_k\mathbf{Z}_0\|_F \leq \|\mathbf{e}_k\|_F \|\mathbf{Z}_0\|_F \leq \sqrt{\delta_k} \cdot \|\mathbf{Z}_0\|_F = \sqrt{\hat{\delta}_k}. \tag{5.59}$$

Therefore, for the conventional case with a fixed coupling matrix, the optimal precoding matrix is obtained by solving $\mathscr{P}_{5.12}$ with $\hat{\delta}_k$ in (5.59), which guarantees that the SINR requirement is met.

## 5.8 Numerical Results

We employ Monte Carlo simulations to evaluate the performance of the proposed precoding methods in this chapter. Since we have assumed $[\mathbf{HW}]_{k,k} > 0, \forall k \in \mathscr{K}$ in (5.16), the obtained precoders can be directly applied to evaluate the BER performance. The system is operating at 2.6GHz and QPSK modulation is applied for the BER evaluation. The simulated channel is generated based on Chapter 2.3.2 and the semi-correlated channel is assumed with $N_p = 50$ DoDs and $\varphi = \pi/8$. We assume a fixed normalised antenna spacing $d = 0.25$ at the transmitter, which is equivalent to $d = \lambda/4$. The initial load impedance vector $\mathbf{z}_L^0$ is $z_{L_i} = 50\Omega, \forall i \in \mathscr{N}$. When quantised load values are considered, we assume $R_{upper} = 100\Omega$, $I_{lower} = -100\Omega$, and $I_{upper} = 100\Omega$ throughout the simulations as representative values, while the tunable range in practical devices can be larger. For the digital SINR balancing optimisations, $\gamma_{min} = 0$ and $\gamma_{max}$ is selected as $\gamma_{\max} = \frac{\max\left(\|\mathbf{h}_k\|^2\right)}{\sigma^2}$. The above parameters remain constant throughout the simulations, unless otherwise stated.

For clarity, the following abbreviations for SINR balancing optimisation are used

- 'SB with MC': FD precoding with a fixed MC effect;
- 'SB no MC': FD precoding assuming no MC effect (as a reference only);

- 'SB A-D Joint': the proposed joint precoding approach with continuous load values;
- 'SB A-D Decoupled': the proposed decoupled method with continuous load values;
- 'SB A-D Sequential': the proposed sequential search algorithm with quantised load values;
- 'SB A-D Exhaustive': the exhaustive search with quantised load values.

As a comparison, we also present the results for the closed-form ZF-based approaches, where we note that the ZF approach is only applicable with

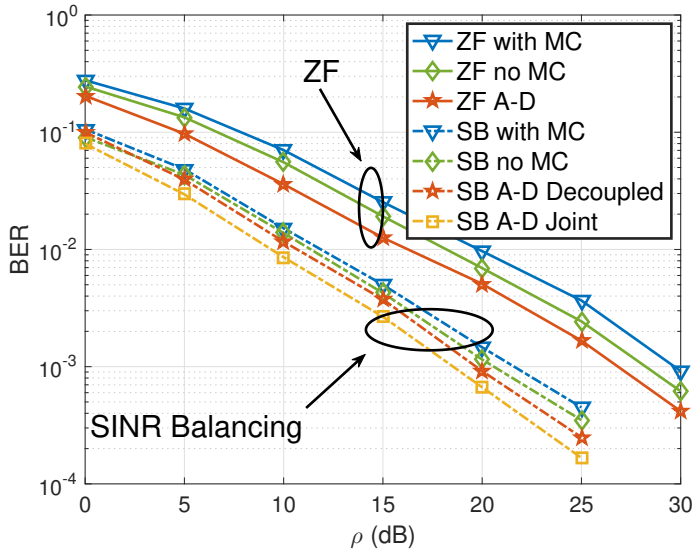


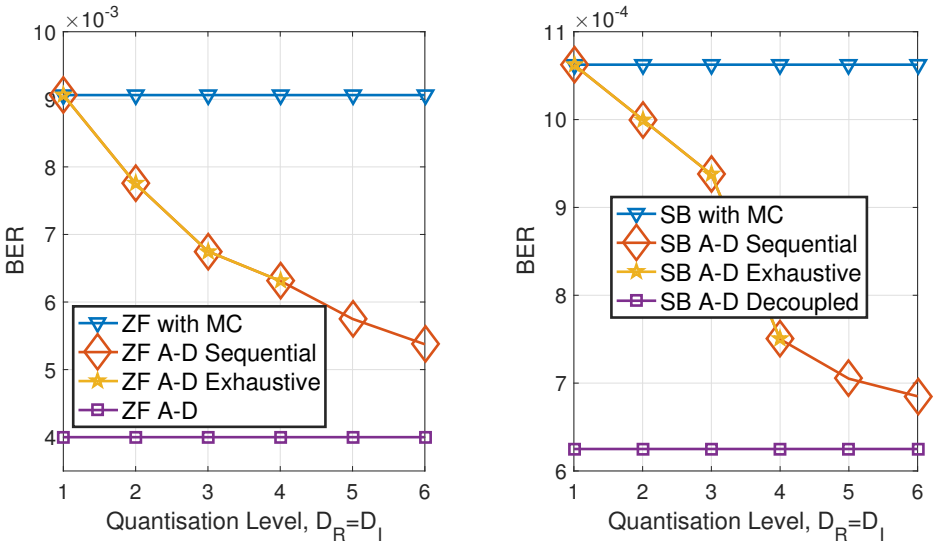

**Figure 5.3:** BER v.s. transmit SNR, $N_t = K = 4$, QPSK, $d = 0.25$, perfect CSI

the decoupled MC exploitation method, and the abbreviations for ZF-based approaches are denoted in a similar way.

### 5.8.1 Perfect CSI

Fig. 5.3 compares the BER performance of the proposed approaches with conventional approaches with fixed coupling effect in perfect CSI. It is observed that for both ZF and SINR balancing, conventional schemes with fixed MC achieve the worst BER performance. For the proposed AD method, the joint precoding achieves the best BER performance with the highest complexity, as the load impedances and the precoders are jointly optimised. For the decoupled approach, while its performance is worse than the joint scheme, it is still superior to the conventional cases with or without a fixed coupling effect. Moreover, in line with existing results, optimisation-based SINR balancing precoding methods achieve an improved performance over ZF.

Fig. 5.4 presents the BER with respect to the quantisation level at SNR=20dB with the employment of quantised antenna loads, where for simplicity we assume $D_R = D_I$. When $D_R = D_I = 1$, the case with quantised loads is identical to the case with fixed MC. For both ZF and SINR balancing pre-




**(a)** ZF **(b)** SINR Balancing

**Figure 5.4:** BER v.s. quantisation level, $N_t = K = 4$, QPSK, $d = 0.25$, perfect CSI, SNR=20dB

coding methods, it is shown that the coupling effect can be beneficial with quantised tunable loads. More importantly, we have observed that for the considered scenario, the proposed sequential search scheme achieves a similar performance to the optimal exhaustive search approach, with a greatly reduced computational cost. The BER result for the exhaustive search is difficult to obtain when the quantisation level is larger than 4, due to the overwhelming computational cost. It is also observed that a larger quantisation level and the resulting larger size of $\mathscr{U}$ leads to an improved performance for the proposed sequential search algorithm.

Since the MC effect is highly relevant to the antenna spacing, in Fig. 5.5 we depict the BER performance with the increasing antenna spacing $d$. As can be observed, with a smaller antenna spacing, the performance gain of the proposed scheme over the case with fixed coupling is more significant, which is due to the stronger coupling effect that can be further exploited. Moreover, for all approaches, we have observed an improved BER with a larger antenna spacing, which is due to the reduced spatial correlation effect.

Fig. 5.6 presents the achievable sum rate for each precoding approach,

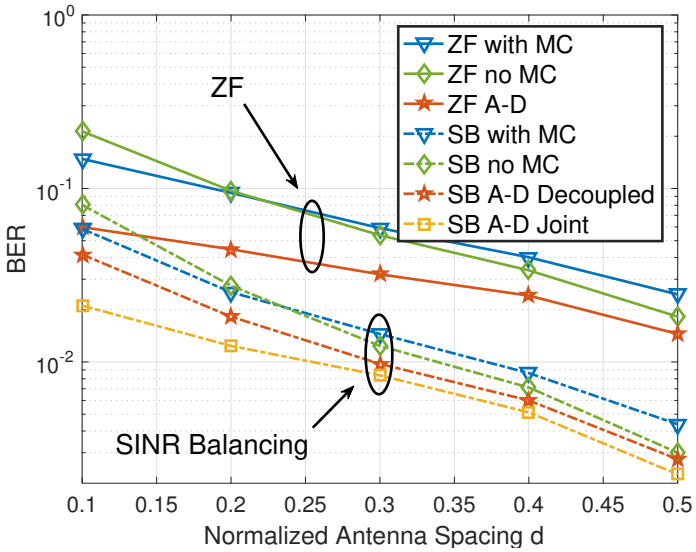



**Figure 5.5:** BER v.s normalised antenna spacing $d$, $N_t = K = 4$, QPSK, perfect CSI

where the channel sum capacity is given as

$$C = \mathbb{E}\left\{ \sup_{\mathbf{G}\in\mathscr{A}} \log_2 \left[ \det\left( \mathbf{I} + \frac{1}{\sigma^2}\mathbf{H}^H\mathbf{G}\mathbf{H} \right) \right] \right\}, \tag{5.60}$$

where sup denotes the supremum function and $\mathscr{A}$ is the set of diagonal $K \times K$ matrices with non-negative elements that ensure $tr\{\mathbf{G}\} = 1$. When uniform transmit power allocation is assumed, $\mathbf{G} = \left(1/K\right) \cdot \mathbf{I}$. It can be seen from

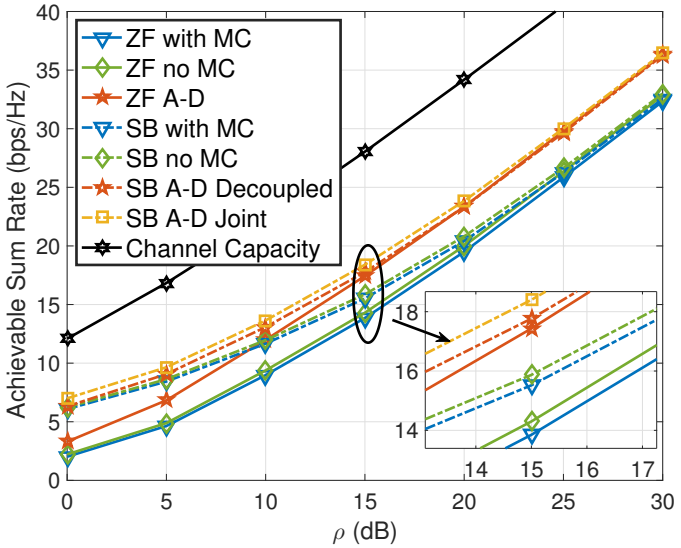



**Figure 5.6:** Sum rate v.s. transmit SNR, $N_t = K = 4$, $d = 0.25$, perfect CSI

Fig. 5.6 that both the joint and the decoupled approach offer a better rate performance compared to conventional digital-only precoding with fixed MC. At the low SNR regime, a rate gain as large as 7 dB can be observed for all SINR balancing based precoding approaches over ZF-based methods.

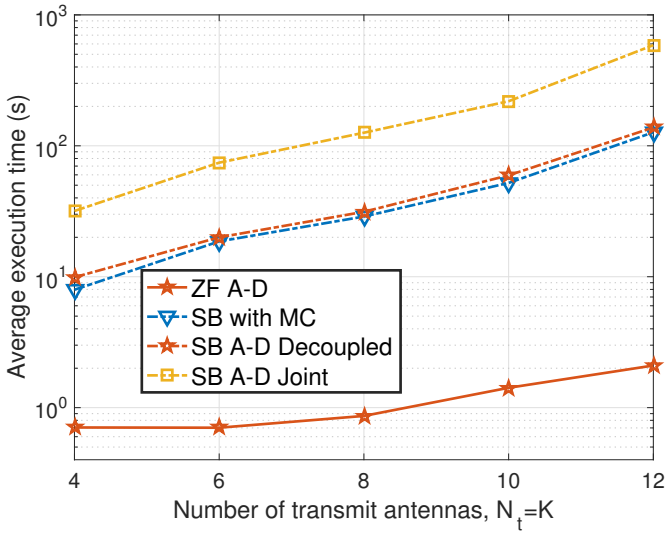
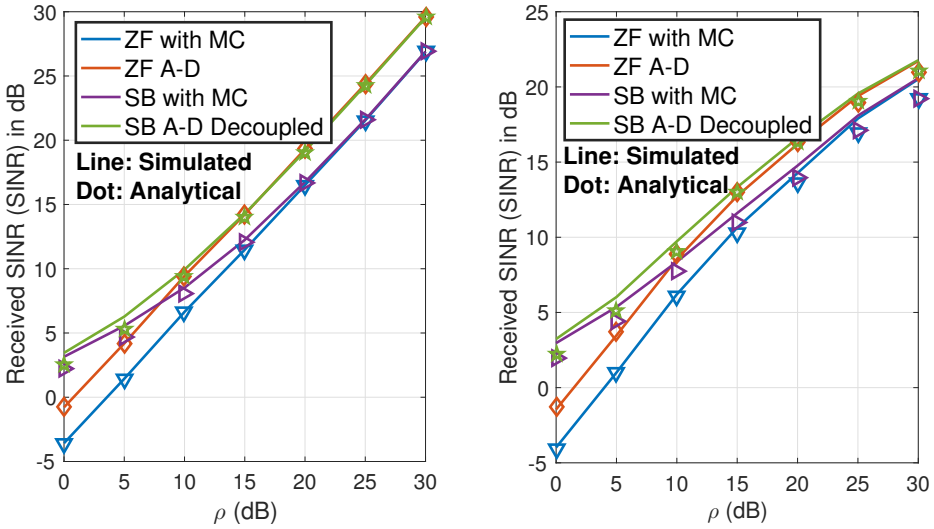



**Figure 5.7:** Average execution time with respect to the number of antennas, $N_t = K$, $N_{max} = 5$

In Fig. 5.7, the computational cost of the proposed joint and the decoupled approaches is numerically illustrated and compared in terms of the average execution time. For the joint scheme, $N_{max} = 5$. We observe that the joint method requires much more execution time than the decoupled approach to obtain the precoder, and this gap becomes increasingly larger with the increase in the problem dimension. Nevertheless, it should be noted that the dominant complexity of the joint scheme is from the SINR balancing procedure instead of the optimisation on the tunable antenna loads, as observed from the negligible gap between 'SB with MC' and 'SB A-D Decoupled'. It is also observed that 'ZF A-D' technique requires the least complexity due to a closed-form precoder in the digital domain.

### 5.8.2 Imperfect CSI - Statistical CSI Errors

Next we present the results for imperfect CSI. We firstly consider the scenarios with statistical CSI errors, where without loss of generality we assume




**(a)** Perfect CSI **(b)** Statistical CSI Errors

**Figure 5.8:** Analytical results v.s. simulated results for ZF and SINR balancing, $N_t = K = 4$, $\alpha = 0.995$, $\beta$=0.1

$\alpha = 0.995$ and $\beta = 0.1$. To validate our analyses on the received SINR, in Fig. 5.8 the analytical results and simulated results of the received SINR with perfect CSI and statistical CSI errors are shown respectively. As can be seen, for both perfect and imperfect CSI cases, a close match between the analytical and simulated results can be observed.

Fig. 5.9 further shows the sum rate performance of the proposed robust approaches with statistical CSI errors. Compared to cases with perfect CSI in Fig. 5.6, all precoding approaches suffer a performance loss due to the channel estimation errors. At the high SNR regime, due to the existence of the delay error matrix $\mathbf{Q}$, the sum rate performance is upper bounded. More importantly, it is observed that the proposed methods that exploit the MC effect still outperform the conventional digital-only precoding methods in the case of imperfect CSI, for both ZF and SINR balancing.

### 5.8.3 Imperfect CSI - Norm-Bounded CSI Errors

We proceed to present the simulation results with norm-bounded CSI errors, where the channel error coefficient is assumed $\delta_k = 0.1, \forall k$. In Fig. 5.10, the

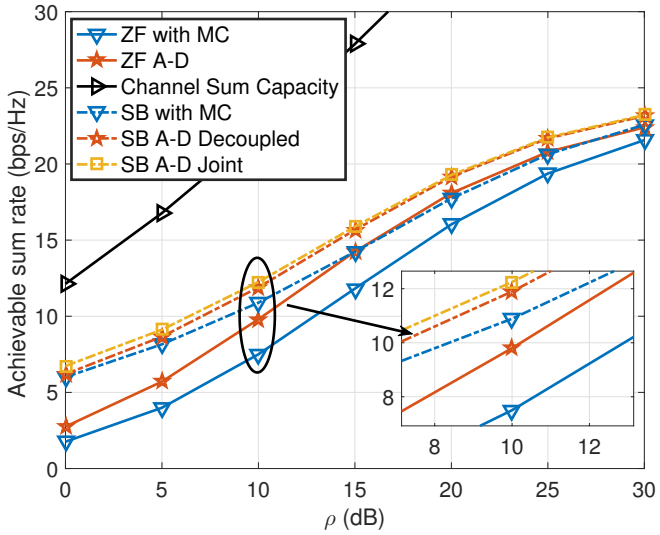



**Figure 5.9:** Sum rate of the robust SINR optimisation v.s. transmit SNR, $N_t = K = 4$, $d = 0.25$, statistical CSI errors, $\alpha = 0.995$, $\beta = 0.1$

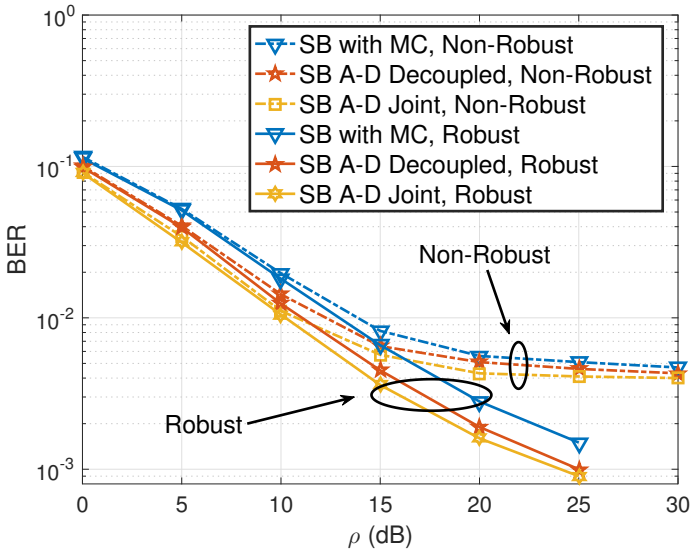



**Figure 5.10:** BER with respect to the transmit SNR, $N_t = K = 4$, QPSK, norm bounded CSI errors, $\delta_k = 0.1$

BER performance is shown with the increasing transmit SNR. Compared to the non-robust approaches that is observed with an error bound at the high SNR regime, the robust scheme achieves an improved performance. Similarly, we observe that the proposed methods that based on the MC exploitation outperform the conventional cases with fixed coupling effect.

## 5.9 Summary

In this chapter, a joint AD optimisation-based precoding that exploits the coupling effect is proposed. We show that by judiciously picking the load values of each antenna element, the MC effect can be beneficial and further improves the system performance. Furthermore, for practical considerations of computational complexity and hardware imperfections, a low-complexity decoupled precoding scheme and a sequential search approach are proposed. In the case of imperfect CSI, the optimality is proven for closed-form precoding approaches, while the robust SINR optimisation methods are further proposed. Simulation results validate the performance improvements by exploiting the MC effect in both perfect and imperfect CSI scenarios.

Both studies in Chapter 4 and Chapter 5 have motivated the use of tunable antenna loads to exploit the MC effect in compact multiple-antenna systems.

# Chapter 6

# Downlink Precoding for ESPARs with Quantised Loads

This chapter is based on our works published in [C4], [C10], [J4].

## 6.1 Introduction

Our previous studies on the coupling effect focus on the generic compact antenna arrays at the BS, where each antenna element is connected to its own RF chain. However, the significant hardware burden imposed by the use of multiple RF chains and the consequent power consumption can be a limitation, especially for the large-scale multiple-antenna systems. Therefore in this chapter, we focus on the downlink transmission design for a power-efficient and hardware-efficient compact parasitic-antenna array structure, known as ESPARs, which has been briefly introduced in Chapter 2.4.1.

As a promising candidate for the future communication systems due to its benefits in hardware complexity and power consumption, ESPARs have received increasing research attention in recent years [8]. Single-fed ESPARs are considered in [78, 79, 131] for spatial multiplexing, where a proof-of-concept experiment has been further demonstrated in [117] to support the practical implementation of parasitic-antenna arrays. Specifically, a loading approach is proposed in [79] to support the multiplexing of two 16QAM signals, in which the effect of impedance errors is numerically studied. Nev-

ertheless, it does not mathematically analyse the impedance error effect on the system performance. In [128], we consider Gaussian impedance errors for ESPAR-based MIMO systems, and the effect of the impedance errors is analytically studied. An important result is reported in [22] that supports an arbitrary precoding for ESPARs by mapping the precoded signals of conventional MIMO arrays on the currents of the ESPAR arrays. In [132], MIMO transmission for parasitic arrays is studied, where the convex optimisation approach is employed to obtain the load values that satisfy the input impedance constraint for the ESPAR array. In [133], a load modulated antenna array is proposed as an extension of the parasitic array, and its application to massive MIMO is discussed. The combination of ESPARs with orthogonal frequency-division multiplexing (OFDM) can be found in [134], where the channel estimation techniques and receiver structures are proposed respectively.

The above existing studies have assumed that each parasitic load has a continuous value range, and can be tuned to any arbitrary values based on the desired radiation pattern. This may not be feasible for practical implementation, since the electronic components implementing the load impedances (varactors, phase shifters, etc) only take values with finite precision, as described in Chapter 5.5. This is particularly critical for the application of parasitic-antenna arrays, as the radiation pattern of the parasitic-antenna arrays is controlled by tuning the load impedances of the passive antennas. Therefore in this chapter, we study practical ESPAR-based MIMO systems with quantised antenna loads, where we focus on precoding techniques at the transmitter side. The system performance for parasitic-antenna arrays in realistic scenarios may be degraded due to the following reasons: i) parasitic load errors due to the quantisation, ii) imperfect CSI due to channel estimation errors. Therefore, we first study the impact of the mismatch effect introduced by the quantised loads and imperfect CSI on the system performance, where we analytically derive the received SINR and probability of

error for ZF precoding. It will be mathematically shown that the impedance errors introduced by the quantisation can be regarded as an additional noise term that is independent of the transmit SNR, which results in an error floor observed in the high SNR regime.

To compensate for the performance degradation introduced by quantised loads, we further propose a quantisation-robust method to approximate the current vector of ESPARs with quantised loads to the ideal current vector by convex optimisation. We propose to jointly optimise the feeding voltage and the quantised loads, where it is further proven that any additional variations in the quantised loads can only lead to an additional performance loss. Therefore, the resulting optimisation problem is reduced to an optimisation on the feeding voltage only, which is convex and can be efficiently solved. Specifically, when single-fed ESPARs are considered, a closed-form solution of the optimal feeding voltage can be obtained, which leads to an efficient employment of the proposed method without incurring significantly additional computational cost. Numerical results show that the proposed quantisation-robust method can compensate for the performance losses and better approach the performance of ideal parasitic-antenna systems, which enables the practical implementation of ESPARs with quantised loads.

For reasons of clarity, we summarise the main contributions of this chapter as:

- For parasitic-antenna systems, we formulate the relationship between the current vector with quantised loads and the ideal current vector, based on which we conduct mathematical analyses on the system performance with ZF precoding, for both perfect CSI and imperfect CSI. It is shown that the quantisation in the loads introduces an additional noise term that is independent of the transmit SNR, which leads to an error floor at high SNRs;

- We propose a quantisation-robust approach to approximate the cur-

rent vector of the practical ESPAR array to the desired signal vector, such that the performance loss introduced by the quantised loads is minimised. It is proven that the optimality is achieved by optimising the feeding voltage only, and the optimal current vector is subsequently obtained via convex optimisation. Specifically, a closed-form solution can be derived for single-fed ESPARs.

## 6.2 System Model

We consider a multi-user scenario in the downlink, where the AP equipped with an ESPAR antenna array communicates with a total number of $K$ single-antenna users simultaneously. Based on the proposition in [22, 132], we consider the current vector of the ESPAR array as the system input, and a general system model that captures the functionality of the parasitic array can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{i} + \mathbf{n}, \tag{6.1}$$

where $\mathbf{y} \in \mathbb{C}^{K\times 1}$ denotes the received signal vector, $\mathbf{H} \in \mathbb{C}^{K\times N_t}$ is the channel matrix from the current vector to the received signals, $\mathbf{i} \in \mathbb{C}^{N_t\times 1}$ denotes the current vector for the parasitic array, $\mathbf{n} \in \mathbb{C}^{K\times 1}$ denotes the noise vector and $\mathbf{n} \sim \mathbb{CN}\left(\mathbf{0}, \sigma^2 \cdot \mathbf{I}\right)$, where $\sigma^2$ is the noise power.

**Channel Model H in ESPARs:** In generic MIMO systems, the transmitting and receiving signals are reflected by the amplitudes and phases of the voltage or current vector. Typically, the channel matrix forms the relationship between the 'input voltages' and the 'output voltages'. When the coupling effect is considered, the relationship between the currents and the feeding voltages is given by (2.58), which is required even in a generic MIMO system. While the general Ohm's law is applicable for both conventional array and parasitic array, conventional MIMO systems typically assume an antenna spacing of larger than half of the wavelength, such that the coupling effect can be negligible in practice. In this case, the mutual impedance matrix becomes a diagonal matrix, which leads to $\mathbf{Z}_m = z_c \cdot \mathbf{I}$, where $z_c$ denotes

the common self-impedance to match the voltage source. The current vector for the conventional MIMO array is obtained as

$$\mathbf{i}_0 = \frac{1}{z_0 + z_c} \cdot \mathbf{v}_0, \tag{6.2}$$

i.e., the currents on each antenna port are a scaled version of the feeding voltages. Accordingly, the channel that relates the 'input currents' and the 'output voltages' can be regarded as a scaled version of the conventional channel matrix, while the statistical properties remain the same. In other words, the channel model developed for conventional MIMO is also valid to relate 'input currents' and 'output voltages'. Therefore, for the ESPAR-based MIMO systems where the input is given by (2.59), current channel models can be directly applied to form the relationship between the 'input currents' and 'output voltages' for parasitic arrays. Accordingly, a semi-correlated Rayleigh channel model described in Chapter 2.3.2 is employed in this chapter.

In the downlink transmission of a $K \times N_t$ MU-MISO system, a generic precoded signal vector can be expressed as

$$\mathbf{x} = \frac{1}{f} \cdot \mathbf{P}\mathbf{s}, \tag{6.3}$$

where $\mathbf{s} \in \mathbb{C}^{K\times 1}$ is data symbol vector with unit power, and $\mathbf{P} \in \mathbb{C}^{N_t \times K}$ denotes the precoding matrix. $f$ is the scaling factor given by

$$f = \|\mathbf{P}\|_F. \tag{6.4}$$

To apply precoding for parasitic-array systems, the precoded signal vector is mapped to the antenna current vector [22, 132], i.e.,

$$\mathbf{i} = \mathbf{x} = \frac{1}{f} \cdot \mathbf{P}\mathbf{s}. \tag{6.5}$$

Then, (6.1) is transformed into

$$\mathbf{y} = \frac{1}{f} \cdot \mathbf{HPs} + \mathbf{n}. \tag{6.6}$$

With this approach, ESPARs can form the same transmit signal and radiate as conventional MIMO, by setting the feeding voltage and the loads as calculated based on (2.62).

## 6.3 Performance Analysis with Quantised Loads and Imperfect CSI

In this chapter, we first introduce the quantisation model in the tunable antenna loads, followed by the performance analyses in both perfect CSI and imperfect CSI.

### 6.3.1 Parasitic Arrays with Quantised Loads

It is observed from (2.62) that the calculation of each tunable load and the feeding voltage is based on the assumption that each load on the parasitic elements can be tuned to any arbitrary continuous values, i.e., infinite-precision tunable loads are assumed. Nevertheless, this is difficult to implement in practice, and in most cases quantised loads of finite precision are employed due to the realistic hardware implementation [135, 136, 137]. Therefore in this chapter, we study the impact of the quantisation in each tunable load on the system performance of ESPARs. We express the value of each quantised load as

$$\hat{z}_m = z_m + e_m^L, \ \forall m \in \mathscr{M}, \tag{6.7}$$

where $\mathscr{M} = \{N+1, N+2, \cdots, N_t\}$. In (6.7), $\hat{z}_m$ denotes the quantised load value at the $m$-th passive antenna element, $z_m$ denotes the ideal load value that forms the desired radiation pattern, and $e_m^L$ denotes the error in the load value due to quantisation. For the quantised loads, we denote $D$ as the quan-

tisation level for both the real part and imaginary part, based on which we obtain the potential values of the quantised loads, given by

$$\hat{z}_m = p_m D + j \cdot q_m D,\ p_m, q_m \in \{0, \pm 1, \pm 2, ...\}. \tag{6.8}$$

Due to the quantisation, the impedance error $e_m^L$ for each tunable load can be regarded as a norm-bounded variable, where the bound is obtained as

$$\left|e_m^L\right|^2 \leq \left(\frac{D}{2}\right)^2 + \left(\frac{D}{2}\right)^2 = \frac{D^2}{2},\ \forall m \in \mathscr{M}. \tag{6.9}$$

### 6.3.2 Analysis - Quantised Loads, Perfect CSI

Before we study the effect of quantisation on the system performance of ESPARs, we first characterise the semi-correlated channel $\mathbf{H}$, where for simplicity we only consider the Rayleigh components $\mathbf{g}_k$. We further assume that a constant steering matrix $\mathbf{A}$ is employed at the BS. Based on (2.47) and the fact that transposition operation does not change the variable distribution, we obtain

$$\mathbf{h}_k^T = (\mathbf{g}_k \mathbf{A})^T = \mathbf{A}^T \mathbf{g}_k^T, \tag{6.10}$$

where $\mathbf{A}^T \in \mathbb{C}^{N_t \times M}$ and $\mathbf{g}_k^T \in \mathbb{C}^{M \times 1}$.

***Lemma***: For a random variable vector $\mathbf{z} \sim \mathbb{CN}(0, \mathbf{K}_z)$, for the linear transformation $\mathbf{y} = \mathbf{B}\mathbf{z}$, we have $\mathbf{y} \sim \mathbb{CN}\left(0, \mathbf{B}\mathbf{K}_z\mathbf{B}^H\right)$ [138].

Based on the lemma above, we obtain that the elements of $\mathbf{h}_k$ also follow a normal distribution with zero mean, where the covariance matrix is given by

$$\mathbf{C}_{\mathbf{h}_k^T} = \mathbf{A}^T \left(\mathbf{A}^T\right)^H = \mathbf{A}^T \cdot conj(\mathbf{A}). \tag{6.11}$$

With perfect CSI, consider each tunable load value with error $e_m^L$ as in (6.7), and the current vector in (2.59) with quantised loads can be rewritten

as

$$\begin{aligned}\hat{\mathbf{i}} &= [diag(\hat{\mathbf{z}}_L) + \mathbf{Z}_M]^{-1}\mathbf{v}_s \\ &= [diag(\mathbf{z}_L) + \mathbf{E}_l + \mathbf{Z}_M]^{-1}\mathbf{v}_s \\ &= (\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{v}_s.\end{aligned} \tag{6.12}$$

In (6.12), $\hat{\mathbf{z}}_L$ is the quantised load vector and we denote

$$\mathbf{E}_L = diag\left\{\left[\mathbf{0}^{1\times N}, e^L_{N+1}, e^L_{N+2}, \cdots, e^L_{N_t}\right]^T\right\} \tag{6.13}$$

as the impedance error matrix, where we note there are no impedance errors for active antenna elements as they employ a fixed load $z_0$. In ideal cases where the parasitic loads with continuous values are employed, the feeding voltages on the active elements can be calculated as (2.62), which remain constant in the presence of impedance errors. This leads to

$$(\mathbf{Z}_T + \mathbf{E}_l)\hat{\mathbf{i}} = \mathbf{Z}_T\mathbf{i}. \tag{6.14}$$

With some further transformations on (6.14), we can formulate the relationship between the current vector with quantised antenna loads and the ideal current vector, given by

$$\begin{aligned}\hat{\mathbf{i}} &= (\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{Z}_T\mathbf{i} \\ &= (\mathbf{Z}_T + \mathbf{E}_l)^{-1}(\mathbf{Z}_T + \mathbf{E}_l - \mathbf{E}_l)\mathbf{i} \\ &= \mathbf{i} - (\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{E}_l\mathbf{i}.\end{aligned} \tag{6.15}$$

Accordingly, the received signal vector in the presence of quantisation errors can be obtained by substituting (6.15) into (6.1), expressed as

$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{H}\hat{\mathbf{i}} + \mathbf{n} \\ &= \mathbf{H}\left[\mathbf{i} - (\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{E}_l\mathbf{i}\right] + \mathbf{n} \\ &= \mathbf{H}\mathbf{i} - \mathbf{H}(\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{E}_l\mathbf{i} + \mathbf{n} \\ &= \mathbf{H}\mathbf{i} + \mathbf{n}_l + \mathbf{n}.\end{aligned} \tag{6.16}$$

Compared to (6.1), the second term $\mathbf{n}_l$ is introduced by the impedance errors due to quantisation, which can be regarded as an additional noise term that is independent of the transmit SNR. We define the equivalent noise term

$$\mathbf{n}_e \triangleq \mathbf{n}_l + \mathbf{n} = -\mathbf{H}(\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{E}_l\mathbf{i} + \mathbf{n}, \tag{6.17}$$

where we note that the errors here are introduced by quantisation in the load impedances, and therefore the AP has the knowledge of the error matrix $\mathbf{E}_l$. Based on (6.17), it is observed that $\mathbf{n}_e$ conditioned on $\mathbf{Z}_T$, $\mathbf{E}_l$ and $\mathbf{i}$ follows an i.i.d. Gaussian distribution with zero mean [34], where the equivalent noise power for each user is given by

$$\upsilon = \omega\zeta^2 + \sigma^2. \tag{6.18}$$

In (6.18), $\omega = \mathrm{tr}\{\mathbf{C}_\mathbf{H}\}$, $\zeta = \left\|(\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{E}_l\right\|_F$ represents the mismatch effect introduced by the quantisation. The current vector is equal to the normalised transmit signal vector, and accordingly we have $\|\mathbf{i}\|_F^2 = 1$.

The above derivation is independent of the precoding methods employed at the AP. To obtain a tractable expression of the received SINR and the resulting analytical probability of error, in the following we assume closed-form ZF precoding at the AP, where based on (6.5) the current vector can be expressed as

$$\mathbf{i} = \frac{1}{f}\cdot\mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\mathbf{s}, \tag{6.19}$$

and the scaling factor $f = \sqrt{tr\left\{(\mathbf{H}\mathbf{H}^H)^{-1}\right\}}$. The received signal vector with quantised loads is obtained as

$$\begin{aligned}\mathbf{y} &= \frac{1}{f}\cdot\mathbf{H}\mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H\right)^{-1}\mathbf{s} + \mathbf{n}_e \\ &= \frac{1}{f}\cdot\mathbf{s} + \mathbf{n}_e,\end{aligned} \tag{6.20}$$

and the resulting received SNR is given by

$$\gamma_k = \frac{1}{f^2\left(\omega\zeta^2 + \sigma^2\right)}. \tag{6.21}$$

It is observed that the first term of the noise is irrelevant to the transmit SNR, and therefore at high transmit SNR regime where $\sigma^2$ becomes negligible, an error floor will be observed.

**Probability of error:** In order to validate our analyses, we introduce the probability of bit error for $\mathbb{N}$-PSK in flat fading with respect to the transmit SNR for $K = N_t$, and the average BER can be calculated as [30, 112, 113] (Appendix C)

$$P_b = \frac{1}{2K}\sum_{k=1}^{K}\left[1 - \sqrt{\frac{\rho_k}{\tau_k^2 + \rho_k}}\right], \tag{6.22}$$

where $\tau_k^2 = \left[\left(\mathbf{C_H}\right)^{-1}\right]_{k,k}$ is the $k$-th diagonal entry in the inverse matrix of the channel covariance $\mathbf{C_H}$. $\rho_k$ denotes the SNR per bit, a.k.a., $\frac{E_b}{N_0}$ for stream $k$. For ZF precoding, $\rho_k$ can be obtained based on $\gamma_k$, expressed as

$$\rho_k = \frac{1}{\log_2\left(\mathbb{N}\right)\cdot K\left(\omega\zeta^2 + \sigma^2\right)},\ \forall k \in \mathscr{K}, \tag{6.23}$$

where $\mathscr{K} = \{1, 2, \cdots, K\}$

### 6.3.3 Analysis - Quantised Loads, Imperfect CSI

We proceed to investigate the performance of ESPAR-based MIMO systems imperfect CSI and quantised loads. We assume the system is operating in the TDD mode, and the imperfect CSI model is given by

$$\mathbf{H} = \hat{\mathbf{H}} + \mathbf{Q}, \tag{6.24}$$

which follows (2.55) described in Chapter 2.3.4, where we have assumed that $\alpha = 1$ for simplicity. To avoid confusion with the impedance error matrix $\mathbf{E}_l$, we employ $\mathbf{Q}$ as the channel error matrix in this section.

With this imperfect CSI model, we substitute (6.15) and (6.24) into (6.1),

which leads to the expression of the received signal vector as

$$\begin{aligned}\mathbf{y} &= \mathbf{H}\left[\mathbf{i} - (\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{E}_l\mathbf{i}\right] + \mathbf{n} \\ &= \left(\hat{\mathbf{H}} + \mathbf{Q}\right)\mathbf{i} - \left(\hat{\mathbf{H}} + \mathbf{Q}\right)(\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{E}_l\mathbf{i} + \mathbf{n} \\ &= \hat{\mathbf{H}}\mathbf{i} + \mathbf{Q}\mathbf{i} - \left(\hat{\mathbf{H}} + \mathbf{Q}\right)(\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{E}_l\mathbf{i} + \mathbf{n},\end{aligned} \tag{6.25}$$

based on which we define an equivalent noise term for imperfect CSI as

$$\hat{\mathbf{n}}_e \triangleq \mathbf{Q}\mathbf{i} - \left(\hat{\mathbf{H}} + \mathbf{Q}\right)(\mathbf{Z}_T + \mathbf{E}_l)^{-1}\mathbf{E}_l\mathbf{i} + \mathbf{n}. \tag{6.26}$$

Compared to the case of perfect CSI, we observe that an additional noise term $\mathbf{Q}\mathbf{i}$ is introduced as a result of the channel estimation errors. Similar to the case of perfect CSI, $\hat{\mathbf{n}}_e$ conditioned on $\mathbf{Q}$, $\mathbf{Z}_T$, $\mathbf{E}_l$ and $\mathbf{i}$ is i.i.d. Gaussian with zero mean, and the equivalent noise power is given by

$$\hat{\upsilon} = \eta + \hat{\omega}\zeta^2 + \sigma^2, \tag{6.27}$$

where $\hat{\omega} = (1+\eta)\,\omega$ considers the channel estimation error effect.

**Probability of error:** The analytical BER for imperfect CSI is also obtained via (6.22). Assuming ZF precoding at the transmitter, we can similarly express the EbN0 for each user $k$ in the case of imperfect CSI as

$$\hat{\rho}_k = \frac{1}{\log_2(\mathbb{N}) \cdot K\left(\eta + \hat{\omega}\zeta^2 + \sigma^2\right)},\ \forall k \in \mathscr{K}. \tag{6.28}$$

## 6.4 Proposed Quantisation-Robust Scheme

As noted above, the presence of quantised loads leads to a performance loss for the ESPAR array at high SNR regimes. Therefore in this chapter, we propose to compensate for this loss by approximating the current vector of practical ESPARs to the desired signal vector with convex optimisation. Specifically, we propose to jointly optimise the feeding voltages and the quantised loads such that an improved performance can be achieved, which is intro-

duced in detail in the following.

### 6.4.1 Multiple-Fed ESPARs

In the presence of quantised load values, we firstly rewrite (6.12) as

$$\hat{\mathbf{i}} = [diag(\hat{\mathbf{z}}_L) + \mathbf{Z}_M]^{-1}\mathbf{v}_s = [diag(\mathbf{z}_L) + \mathbf{Z}_M + \mathbf{E}_l]^{-1}\mathbf{v}_s, \tag{6.29}$$

where $\hat{\mathbf{z}}_L$ denotes the quantised load vector. $\mathbf{E}_l$ is given in (6.13), where we note that there exist quantisation errors only for the tunable loads. Different from previous chapter where we fix the values of the feeding voltages, we consider the variations in both the quantised loads and the voltages, which leads to the expression of the current vector as

$$\mathbf{i}^R = \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_M\right]^{-1}\mathbf{v}_s^R = [diag(\mathbf{z}_L) + \mathbf{Z}_M + \mathbf{E}_l + D \cdot diag(\mathbf{t})]^{-1}(\mathbf{v}_s + \Delta\mathbf{v}), \tag{6.30}$$

where $\mathbf{i}^R$ denotes the optimised current vector, the optimised quantised loads are denoted as $\hat{\mathbf{z}}_L^R$, $\mathbf{v}_s^R$ denotes the optimised voltage vector and $\Delta\mathbf{v}$ is the variation in the feeding voltages. In (6.30), $\mathbf{E}^R = \mathbf{E} + D \cdot diag(\mathbf{t})$, $\mathbf{t} \in \mathbb{CZ}^{N_t \times 1}$ is the complex integer vector to be optimised, which satisfies $\mathbf{t}(n) = 0$, $n \in \mathscr{N}$, where $\mathscr{N} = \{1, 2, \cdots, N\}$. $D \cdot diag(\mathbf{t})$ represents the variations in the quantised load values when the optimality is reached. Based on (2.59), (6.30) can be further transformed into

$$\begin{aligned}
\mathbf{i}^R &= \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_M\right]^{-1}\{[diag(\hat{\mathbf{z}}_L) + \mathbf{Z}_M]\mathbf{i} + \Delta\mathbf{v}\} \\
&= \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_M\right]^{-1}\left\{\left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_M - \mathbf{E}_l^R\right]\mathbf{i} + \Delta\mathbf{v}\right\} \\
&= \mathbf{i} - \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_M\right]^{-1}\mathbf{E}_l^R\mathbf{i} + \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_M\right]^{-1}\Delta\mathbf{v} \\
&= \mathbf{i} + \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_M\right]^{-1}\left(\Delta\mathbf{v} - \mathbf{E}_l^R\mathbf{i}\right).
\end{aligned} \tag{6.31}$$

Based on the above, we can express the difference between the desired current vector and the optimised current vector as

$$\Delta\mathbf{i} = \mathbf{i}^R - \mathbf{i} = \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_m\right]^{-1}\left(\Delta\mathbf{v} - \mathbf{E}^R\mathbf{i}\right), \tag{6.32}$$

which leads to the following proposition.

**Proposition 1:** $\Delta\mathbf{i}$ cannot be minimised to $\mathbf{0}$ by optimising the feeding voltage vector or the quantised load values.

**Proof:** $\Delta\mathbf{i} = \mathbf{0}$ is equivalent to

$$\left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_m\right]^{-1}\left(\Delta\mathbf{v} - \mathbf{E}^R\mathbf{i}\right) = \mathbf{0}. \tag{6.33}$$

Since $\Delta\mathbf{v}(m) = 0$, $\forall m \in \mathscr{M}$, we obtain that $\Delta\mathbf{v} - \mathbf{E}^R\mathbf{i} \neq \mathbf{0}$. Accordingly, (6.33) is equivalent to finding non-zero solutions for a linear system $\mathbf{U}\mathbf{x} = \mathbf{0}$ with $\mathbf{U} = \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_m\right]^{-1}$. Based on the linear algebra theory, the following condition must be satisfied

$$\det\left\{\left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_m\right]^{-1}\right\} = 0. \tag{6.34}$$

However, (6.34) is not achievable for an inverse matrix, which completes the proof. ■

**Proposition 1** implies that there always exists a performance loss compared to the conventional FD MIMO systems, when quantised loads are employed for parasitic arrays. In the following, we propose to minimise the performance loss by optimising the values of both the quantised loads and the voltages, where the following proposition is given.

**Proposition 2:** When the optimality is achieved, $\mathbf{t}^* = \mathbf{0}$, which means that any further variations in the quantised loads will only incur additional performance losses.

**Proof:** When a variation in the quantised loads is further introduced, $\mathbf{t} \neq \mathbf{0}$ and based on (6.33) we can further obtain

$$\mathbf{i}^R = \mathbf{i} - \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_M\right]^{-1}\mathbf{E}_l\mathbf{i} + \left[diag\left(\hat{\mathbf{z}}_L^R\right) + \mathbf{Z}_M\right]^{-1}\left[\Delta\mathbf{v} - D\cdot diag\left(\mathbf{t}\right)\mathbf{i}\right]. \tag{6.35}$$

Based on the expression of $\Delta\mathbf{v}$ and $\mathbf{t}$, it is observed that the introduction of $\mathbf{t}$ can only increase the norm of $[\Delta\mathbf{v} - D\cdot diag\left(\mathbf{t}\right)\mathbf{i}]$, which further leads to an

increased deviation from the desired current vector $\mathbf{i}$. Moreover, the deviation on the current values in the parasitic elements cannot be compensated by the voltage vector. Therefore, to keep the deviation as small as possible, the optimal case is to keep the load values unchanged and only optimise the voltages only. ■

Based on the above two propositions, we can construct the optimisation problem as

$$\begin{aligned} \mathscr{P}_{6.1}: \ & \min_{\Delta\mathbf{v}} \ \|\Delta\mathbf{i}\|_F^2 \\ & s.t. \quad \Delta\mathbf{v}(m) = 0, \ \forall m \in \mathscr{M} \end{aligned} \tag{6.36}$$

By introducing

$$\mathbf{P} = [diag(\mathbf{z}_L) + \mathbf{Z}_M + \mathbf{E}_l] \tag{6.37}$$

which is fixed with respect to $\Delta\mathbf{v}$, $\mathscr{P}_{6.1}$ can be transformed into a convex form as

$$\begin{aligned} \mathscr{P}_{6.2}: \ & \min_{\Delta\mathbf{v}} \ \left\|\mathbf{P}^{-1}\Delta\mathbf{v} - \mathbf{P}^{-1}\mathbf{E}\mathbf{i}\right\|_F^2 \\ & s.t. \quad \Delta\mathbf{v}(m) = 0, \ \forall m \in \mathscr{M} \end{aligned} \tag{6.38}$$

which can be efficiently solved by convex optimisation tools.

### 6.4.2 Single-Fed ESPARs: A Closed-Form Solution

We further consider a special case where single-fed ESPARs are employed, and show that the optimal solution can be obtained in a closed form. In the case of single-fed ESPARs, $N = 1$, and by introducing an auxiliary complex variable $\alpha$, $\Delta\mathbf{v}$ can be expressed as

$$\Delta\mathbf{v} = \alpha \cdot \mathbf{v}_s, \tag{6.39}$$

as there is only one entry in $\mathbf{v}_s$ that is non-zero in the case of single-fed ESPARs. Based on (6.32) and the expression of $\mathbf{P}$ in (6.37), the difference in the

current vector $\Delta\mathbf{i}$ for single-fed ESPARs can be further simplified into

$$\begin{aligned}\Delta\mathbf{i} &= (1+\alpha)\,\mathbf{P}^{-1}(\mathbf{P}-\mathbf{E})\,\mathbf{i}-\mathbf{i}\\ &= \left(\mathbf{I}-\mathbf{P}^{-1}\mathbf{E}\right)\mathbf{i}\cdot\alpha+\left[\left(\mathbf{I}-\mathbf{P}^{-1}\mathbf{E}\right)-\mathbf{I}\right]\mathbf{i}\\ &= \mathbf{T}\cdot\alpha+(\mathbf{T}-\mathbf{i}),\end{aligned} \tag{6.40}$$

where the expression of $\mathbf{T}$ is given by

$$\mathbf{T} = \left(\mathbf{I}-\mathbf{P}^{-1}\mathbf{E}\right)\mathbf{i} \tag{6.41}$$

with $\mathbf{T}\in\mathbb{C}^{N_t\times 1}$. With the above transformations, $\|\Delta\mathbf{i}\|^2$ can be further expressed as

$$\begin{aligned}\|\Delta\mathbf{i}\|_F^2 &= tr\left\{\Delta\mathbf{i}\Delta\mathbf{i}^H\right\}\\ &= tr\left\{\left[\mathbf{T}\cdot\alpha+(\mathbf{T}-\mathbf{i})\right]\left[\mathbf{T}\cdot\alpha+(\mathbf{T}-\mathbf{i})\right]^H\right\}\\ &= tr\left\{\mathbf{T}^H\mathbf{T}\cdot\alpha\alpha^*\right\}+2\Re\left(tr\left\{\left(\mathbf{T}\mathbf{T}^H-\mathbf{T}\mathbf{i}^H\right)\cdot\alpha\right\}\right)+tr\left\{(\mathbf{T}-\mathbf{i})(\mathbf{T}-\mathbf{i})^H\right\}.\end{aligned} \tag{6.42}$$

Since $\mathbf{T}$ and $\mathbf{i}$ are independent of $\alpha$, the optimisation on $\|\Delta\mathbf{i}\|_F^2$ for single-fed ESPAR arrays can be constructed as

$$\mathscr{P}_{6.3}:\ \min_{\alpha}\ f(\alpha) \tag{6.43}$$

where the objective function $f(\alpha)$ is given by

$$\begin{aligned}f(\alpha) &= tr\left\{\mathbf{T}^H\mathbf{T}\alpha\alpha^*\right\}+2\Re\left(tr\left\{\left(\mathbf{T}\mathbf{T}^H-\mathbf{T}\mathbf{i}^H\right)\alpha\right\}\right)\\ &= tr\left\{S\cdot\alpha\alpha^*\right\}+2\Re\left(tr\left\{\mathbf{T}\cdot\alpha\right\}\right).\end{aligned} \tag{6.44}$$

In (6.44), $S$ and $\mathbf{U}$ are given by

$$S=\mathbf{T}^H\mathbf{T}>0,\ \mathbf{U}=\mathbf{T}\mathbf{T}^H-\mathbf{T}\mathbf{i}^H. \tag{6.45}$$

By further denoting

$$\Re(\alpha) = \alpha_{RE},\ \Im(\alpha) = \alpha_{IM}, \tag{6.46}$$

the objective function in $\mathscr{P}_{6.3}$ can be further obtained as

$$f(\alpha) = \left\{ S\alpha_{RE}^2 + 2\sum_{i=1}^{N_t} \Re\left[\mathbf{U}(i,i)\right]\alpha_{RE} \right\} + \left\{ S\alpha_{IM}^2 - 2\sum_{i=1}^{N_t} \Im\left[\mathbf{U}(i,i)\right]\alpha_{IM} \right\}. \tag{6.47}$$

It can be observed that the real part and imaginary part of $f(\alpha)$ are both in a quadratic form. Since $\alpha$ is a scalar, we can obtain the optimal $\alpha^*$ as

$$\alpha^* = \frac{-\sum_{i=1}^{N_t} \Re\left[\mathbf{U}(i,i)\right]}{S} + j \cdot \frac{\sum_{i=1}^{N_t} \Im\left[\mathbf{U}(i,i)\right]}{S}. \tag{6.48}$$

Accordingly, the optimal feeding voltage for single-fed ESPARs is obtained as

$$\mathbf{v}_s^R = (1+\alpha^*)\,\mathbf{v}_s. \tag{6.49}$$

## 6.5 Numerical Results

To evaluate the performance of ESPAR-based MIMO systems and the proposed quantisation-robust approaches, in this section we present numerical results based on Monte Carlo simulations. QPSK modulation is employed to evaluate the BER performance. For the parasitic array, we assume it operates at the frequency of 2.6GHz and the antenna spacing is $d = \lambda/4$. For the semi-correlated channels described in Chapter 2.3.2, we assume $M = 50$ throughout the simulations. We consider both multi-fed and single-fed parasitic antenna arrays, where the transmit SNR is defined as $\rho = \frac{1}{\sigma^2}$. Closed-form ZF precoding is assumed throughout the simulations.

### 6.5.1 Single-Fed Parasitic Arrays

We first focus on the small-scale single-fed parasitic arrays. To validate our analysis in Chapter 6.3, Fig. 6.1 compares the analytical probability of bit error and the numerical results for ZF precoding with impedance errors in

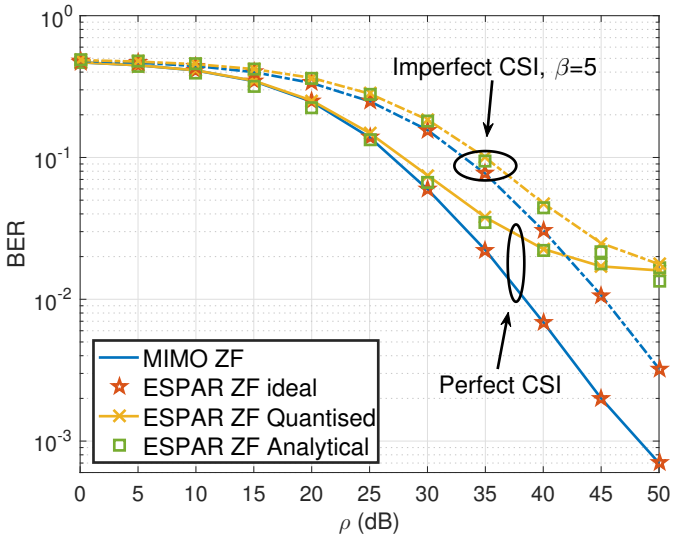



**Figure 6.1:** BER v.s. transmit SNR, single-fed ESPARs, $N_t = K = 5$, $D = 1$, QPSK, $\beta = 5$ for imperfect CSI

both perfect CSI and imperfect CSI, where the channel error coefficient is assumed to be $\beta = 5$ for imperfect CSI. It can be seen that ideal ESPAR-based MIMO systems with a single RF chain can achieve the same performance as conventional MIMO with full RF chains. When we employ practical quantised loads, a severe performance degradation is observed and an error floor appears at high SNR regimes. For both perfect CSI and imperfect CSI, it is observed that the analytical BER results match the simulated results. Note that the corresponding curves for perfect and imperfect CSI converge to the same error floor, which signifies that the quantisation errors become dominant at high SNR.

Fig. 6.2 presents the BER of the proposed quantisation-robust scheme for single-fed ESPAR systems with $N_t = 5$ and $K = 3$, in both perfect and imperfect CSI with the increasing transmit SNR. With the optimised feeding voltage obtained by the proposed closed-form method, the performance of ESPAR arrays with quantised loads are significantly improved, both for perfect and imperfect CSI. The proposed approach therefore enables the use of practical parasitic antenna arrays with the employment of quantised antenna loads.

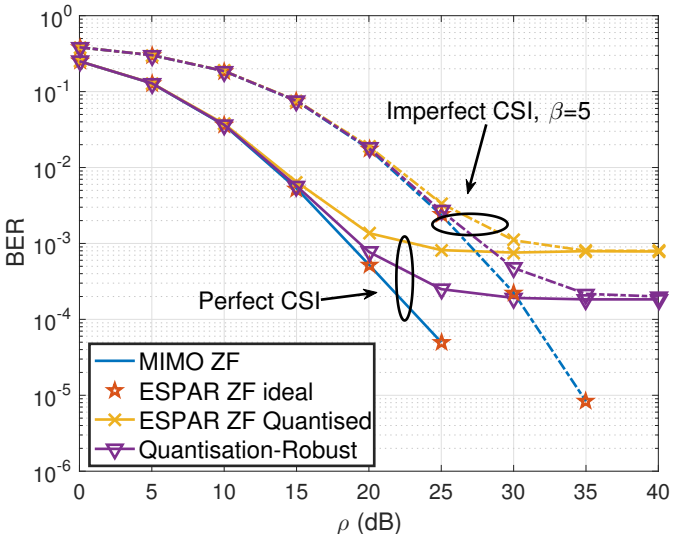



**Figure 6.2:** BER v.s. transmit SNR for the proposed quantisation-robust scheme, single-fed ESPARs, $N_t = 5$, $K = 3$, $D = 1$, QPSK, $\beta = 5$ for imperfect CSI

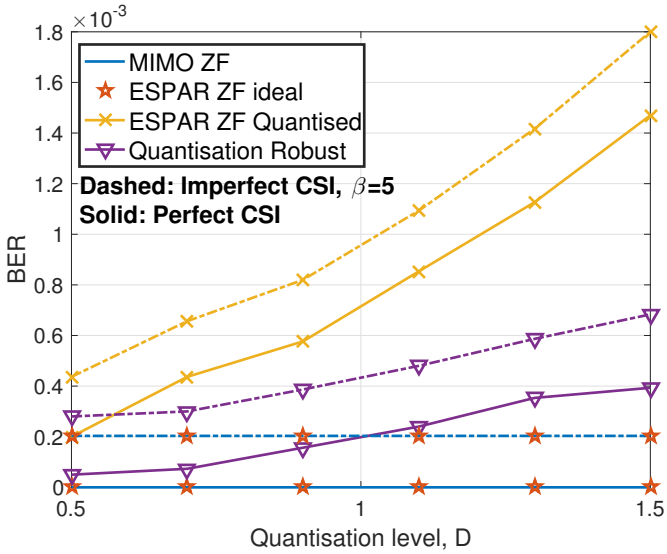



**Figure 6.3:** BER v.s. quantisation level $D$, single-fed ESPARs, $N_t = 5$, $K = 3$, SNR=30dB, QPSK, $\beta = 5$ for imperfect CSI

Fig. 6.3 compares the BER performance of conventional MIMO and ESPARs with an increasing quantisation level $D$ at the SNR of 30dB, in both perfect and imperfect CSI. The BER performance for conventional MIMO systems remains constant, as fixed antenna loads $z_0$ are employed for each antenna element. For parasitic arrays with quantised loads, the BER performance is degraded severely with the increase in the quantisation level.

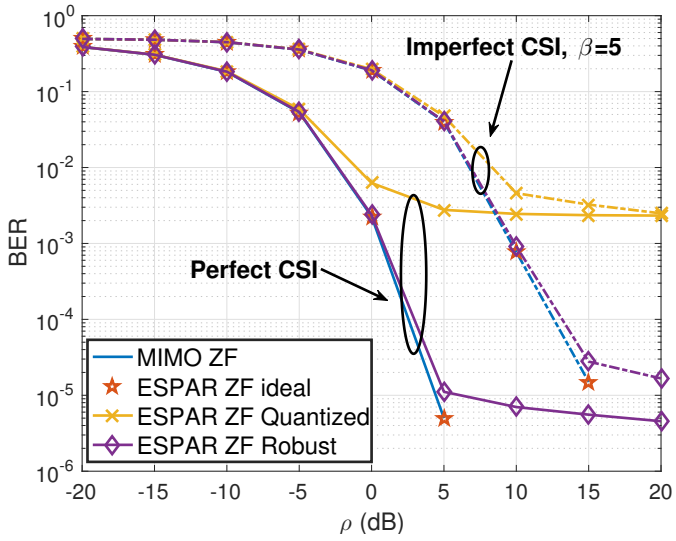
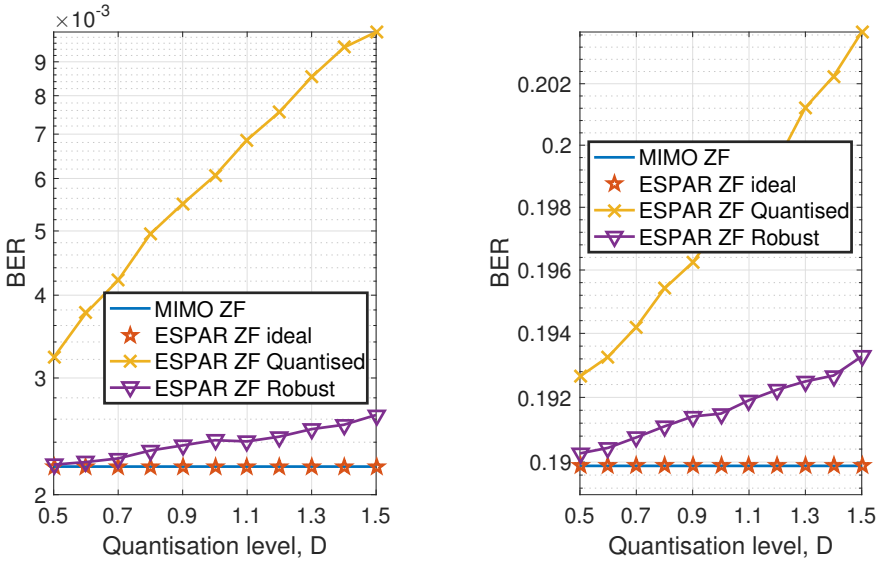



**Figure 6.4:** BER v.s. transmit SNR, multiple-fed ESPARs, $N_t = 64$, $K = 6$, $N = 7$, $D = 1$, QPSK, $\beta = 5$ for imperfect CSI




**(a)** Perfect CSI **(b)** Imperfect CSI, $\beta = 5$

**Figure 6.5:** BER v.s. quantisation level $D$, multiple-fed ESPARs, $N_t = 64$, $K = 6$, $N = 7$, SNR=0dB, QPSK

For the proposed quantisation-robust approach, it can well alleviate the impedance mismatch effect and achieve an improved performance, in both perfect and imperfect CSI. It is also observed that the performance gains of the proposed approach over ESPARs with quantised loads become more significant with the increase in the quantisation level.

### 6.5.2 Multiple-Fed Parasitic Arrays

We proceed to present the results for multiple-fed ESPARs. In Fig. 6.4, the BER performance of multiple-fed ESPARs is shown with respect to the increasing transmit SNR, where it is observed that the ideal multi-fed ESPAR systems can achieve a similar performance to the conventional MIMO systems, while the presence of quantisation in the load values severely degrades the detection performance. Specifically for the multiple-fed ESPARs, the proposed quantisation-robust approach can greatly alleviate the performance loss by quantisation and achieve a close-to-optimal performance. In terms of the BER performance with respect to the quantisation level, a similar result to single-fed ESPARs can be observed in Fig. 6.5, where it is observed that the proposed quantisation-robust scheme is less sensitive to the increase in the quantisation interval, which validates its robustness.

## 6.6 Summary

In this chapter, the precoding techniques for parasitic antenna arrays with quantised loads are studied. Our mathematical analyses indicate that impedance errors introduced by quantisation introduce an additional noise term, which is validated numerically. Furthermore, we propose a quantisation-robust scheme to compensate for the performance loss by quantisation, where the optimal feeding voltage for single-fed ESPARs can be obtained in a closed form. Numerical results show that ideal ESPARs can achieve a similar performance to the conventional MIMO systems, while the proposed quantisation-robust method can well alleviate the performance degradation by the use of quantised loads.

# Chapter 7

# Hybrid Precoding and Application to SWIPT

This chapter is based on our works in [C6], [C8], [J2], [J6].

## 7.1 Hybrid Precoding with Virtual Path Selection

### 7.1.1 Introduction

In addition to the hybrid structures considered in Chapter 6 with parasitic antenna elements, in this chapter we extend our study to hybrid structures with phase shifters, introduced in Chapter 2.4.2.

With the rich and under-exploited spectrum resources in the mmWave band, mmWave communications have been one of the promising candidates for the future 5G communication systems [65, 139, 140]. Compared to existing microwave cellular systems, the operating frequency for mmWave communications is in tens of GHz, which results in a much higher pathloss in the free space [65]. To overcome the pathloss encountered in the mmWave frequencies, mmWave communication systems will require a large-scale antenna array at the BS [66]. Thanks to the high frequency, the consequent small carrier wavelength for mmWave enables packing a large number of antennas in a small physical space [140]. Nevertheless, the conventional FD precoding techniques require a dedicated RF chain for each antenna element, which becomes impractical in the large-scale MIMO systems due to

the significant cost and power consumption in the hardware components, especially for mixed-signal hardwares designed in the mmWave band.

At first, analog processing and RF precoding techniques have been designed for mmWave systems [141, 142], where low-cost phase shifters are employed to circumvent the above practical constraints. However, there exists a performance gap compared to the FD precoding methods, especially for multi-user transmission. This is due to the limited capability in managing interference for analog-only precoding techniques. More recently, to enable spatial multiplexing and approach the performance of FD precoding without incurring significant hardware costs, the hybrid AD precoding has been introduced as a promising technique to balance the performance and the cost [11, 23]. The hybrid precoding employs a limited number of RF chains and consists of an analog precoding via the phase shifter networks and a low-dimensional digital precoding that manages the multi-stream and multi-user interference. Thanks to the reduced number of RF chains, the hybrid structure is also shown to achieve an improved power efficiency compared to the FD systems, which meets the energy-efficient transmission for the future wireless communication systems. The power efficiency gains of the hybrid precoding will be studied in the second part of this chapter for SWIPT.

Hybrid precoding techniques have received increasing research attention due to the benefits mentioned above [67, 80, 143]. In [80], hybrid designs for single-user mmWave MIMO systems have been proposed, where the orthogonal match pursuit (OMP) algorithm is proposed to obtain the analog and baseband precoding matrices. The extension to multi-user cases has been studied in [67], where a codebook-based design is proposed. Nevertheless, we note that the above designs for mmWave systems assume a fully-connected RF structure, where each antenna is connected to all RF chains. More recent study in [143] has shown that the fully-connected hybrid structure will introduce significant insertion losses, which degrade the perfor-

mance in practical mmWave systems. Subsequently, a partially-connected structure becomes more promising for practical implementation [81], where each antenna is only connected to one RF chain and a lower number of phase shifters is required.

In this chapter, we firstly consider the multi-user transmission in the downlink mmWave communication systems, and propose a low-complexity hybrid precoding design. We perform a virtual path selection in the RF domain to maximise the effective analog channel gain, and we subsequently employ ZF precoding in the baseband to eliminate the multi-user interference, based on the low-dimensional effective analog channel. Three distinct designs are proposed respectively with each achieving a different performance-complexity tradeoff. The proposed precoding method applies to both the fully-connected and partially-connected structures, which can approach the performance of FD precoding schemes with a reduced hardware cost. Our computational cost analysis and simulation results validate the performance-complexity advantages of the proposed schemes over existing techniques.

### 7.1.2 System Model

We consider a large-scale MU-MIMO downlink system that follows Chapter 2.4.2, where a BS with $N_t$ antennas and $N_{RF}^t$ RF chains communicates with $K$ users, each equipped with $N_r^k$ antennas. Following [144], we focus on the case where the BS employs $K$ RF chains to support the transmission of $K$ streams, i.e. $K = N_{RF}^t$, and single-stream transmission is assumed for each user. With the hybrid AD structures, the transmit signal vector can be constructed as

$$\mathbf{x} = \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s}, \tag{7.1}$$

where $\mathbf{F}_{BB} = [\mathbf{f}_{BB_1}, \mathbf{f}_{BB_2}, \cdots, \mathbf{f}_{BB_K}] \in \mathbb{C}^{K\times K}$ is the baseband precoder. $\mathbf{s} \in \mathbb{C}^{K\times 1}$ denotes the data symbol vector and $\mathbb{E}\left\{\mathbf{s}\mathbf{s}^H\right\} = \frac{P}{K}\cdot\mathbf{I}$, where $P$ is the total available transmit power and we assume uniform power allocation for each

stream. $\mathbf{F}_{RF} \in \mathbb{C}^{N_t \times K}$ is the RF precoder implemented with phase shifters, and therefore each entry of $\mathbf{F}_{RF}$ is of constant modulus. The total power constraint is satisfied by normalising the digital precoder $\mathbf{F}_{BB}$, such that

$$\|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 = K \tag{7.2}$$

is obtained [67]. At the receiver side, each user $k$ is equipped with one RF chain, and a RF-only combiner $\mathbf{w}_k^H \in \mathbb{C}^{1 \times N_r^k}$ is employed, where each entry in the vector $\mathbf{w}_k$ is of constant modulus and normalised to satisfy

$$[\mathbf{w}_k]_m = \frac{1}{\sqrt{N_r^k}},\ \forall k \in \mathscr{K},\ \forall m \in \left\{1, 2, \cdots, N_r^k\right\}, \tag{7.3}$$

where $\mathscr{K} = \{1, 2, \cdots, K\}$. The received symbol for user $k$ can be expressed as

$$r_k = \mathbf{w}_k^H \mathbf{H}_k \mathbf{F}_{RF} \mathbf{f}_{BB_k} s_k + \mathbf{w}_k^H \mathbf{H}_k \mathbf{F}_{RF} \sum_{j \neq k} \mathbf{f}_{BB_j} s_j + \mathbf{w}_k^H \mathbf{n}_k, \tag{7.4}$$

where $\mathbf{H}_k \in \mathbb{C}^{N_r^k \times N_t}$ is the mmWave channel from the BS to user $k$, which is introduced in Chapter 2.3.3, $\mathbf{n}_k$ is the additive Gaussian noise vector with zero mean and covariance $\sigma^2 \cdot \mathbf{I}$. The sum rate is then calculated according to [67]

$$R = \sum_{k=1}^{K} \log_2 \left( 1 + \frac{\frac{P}{K} \left| \mathbf{w}_k^H \mathbf{H}_k \mathbf{F}_{RF} \mathbf{f}_{BB_k} \right|^2}{\sum\limits_{j \neq k}^{K} \frac{P}{K} \left| \mathbf{w}_k^H \mathbf{H}_k \mathbf{F}_{RF} \mathbf{f}_{BB_j} \right|^2 + \sigma^2} \right). \tag{7.5}$$

### 7.1.3 Proposed Hybrid Precoding Approaches

We introduce the proposed hybrid precoding techniques in this chapter, where the analog domain is first designed by applying phase-only weights, followed by the design in the digital domain. We first consider the fully-connected structures, and then extend to the partially-connected structures. By denoting

$$\mathbf{W} = diag\left\{[\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_K]\right\},\ \mathbf{H} = \left[\mathbf{H}_1^T, \mathbf{H}_2^T, \cdots, \mathbf{H}_K^T\right]^T, \tag{7.6}$$

we can express the effective RF channel in a compact form as

$$\tilde{\mathbf{H}} = \mathbf{W}^H \mathbf{H} \mathbf{F}_{RF}, \tag{7.7}$$

where we note that the diagonal entries in $\tilde{\mathbf{H}}$ are the RF-to-RF channel gains, while the off-diagonal terms denote interference. Therefore, we focus on the maximisation on the diagonal terms of $\tilde{\mathbf{H}}$ in the analog domain, while the off-diagonal interference is handled by the low-dimensional baseband precoder. Based on (7.7), the $k$-th diagonal entry of $\tilde{\mathbf{H}}$ can be expressed as

$$\tilde{\mathbf{H}}(k,k) = \mathbf{w}_k^H \mathbf{H}_k \mathbf{f}_k^f. \tag{7.8}$$

The aim of the proposed approaches is to extract the phases of the downlink channel from the BS to users to formulate the analog precoder and combiner, where we employ the conjugate transpose to maximise the analog effective channel gain. We note that in a MU-MIMO system, for each user $k$, $\mathbf{H}_k \in \mathbb{C}^{N_r^k \times N_t}$, while $\mathbf{f}_k^f \in \mathbb{C}^{N_t \times 1}$. Therefore, direct conjugate transposition cannot be applied. Nevertheless, as $\mathbf{H}_k$ is $N_r^k \times N_t$ dimensional, we decompose it into

$$\mathbf{H}_k = \left[ \left(\mathbf{h}_k^1\right)^T, \left(\mathbf{h}_k^2\right)^T, ..., \left(\mathbf{h}_k^{N_r^k}\right)^T \right]^T, \tag{7.9}$$

and each $\mathbf{h}_k^i$ represents a virtual path. Accordingly, in the following we propose three distinct precoding approaches based on virtual path selection to maximise $\tilde{\mathbf{H}}(k,k)$ for each user $k$.

**Joint Design:** In the joint design, we propose to jointly design $\mathbf{w}_k$ and $\mathbf{f}_k^f$ for each user $k$. In the first step, for each $i \in \left\{1, 2, ..., N_r^k\right\}$, we design a temporary analog precoder $\mathbf{p}_k^i$ as the conjugate transpose of $\mathbf{h}_k^i$ with an element-wise normalisation to satisfy the constant modulus constraint, given by [144]

$$\left[\mathbf{p}_k^i\right]_m = \frac{1}{\sqrt{N_t}} e^{j \cdot \left(-\varphi_m^i\right)}, \; m \in \{1, 2, ..., N_t\}, \tag{7.10}$$

where $\varphi_m^i$ is the phase of the $m$-th element in $\mathbf{h}_k^i$, and we further denote

$$\mathbf{t}_k^i = \mathbf{H}_k \mathbf{p}_k^i. \tag{7.11}$$

Similarly, a temporary analog combiner $\mathbf{q}_k$ is constructed as

$$\left[\mathbf{q}_k^i\right]_m = \frac{1}{\sqrt{N_r^k}} e^{j\cdot\omega_m^i},\ m \in \left\{1,2,...,N_r^k\right\}, \tag{7.12}$$

where $\omega_m^i$ denotes the phase of the $m$th element in $\mathbf{t}_k^i$. In the second step, we select the virtual path $i$ that achieves the largest effective channel gain $\tilde{\mathbf{H}}(k,k)$, expressed as

$$\left(i^*, \mathbf{p}_k^{i^*}, \mathbf{q}_k^{i^*}\right) = \underset{i\in\left\{1,2,...,N_r^k\right\}}{\arg\max} \left[\left(\mathbf{q}_k^i\right)^H \mathbf{H}_k \mathbf{p}_k^i\right]. \tag{7.13}$$

The detailed joint approach is given in Algorithm 7.1.

**Algorithm 7.1** Joint Analog Precoder and Combiner Design

**input** : $\mathbf{H}_k$
**output** : $\mathbf{f}_k^f$, $\mathbf{w}_k$
**for** $i = 1$ to $N_r^k$ **do**
  Calculate $\left[\mathbf{p}_k^i\right]_m = \frac{1}{\sqrt{N_t}} e^{j\cdot\left(-\varphi_m^i\right)}, m \in \{1,2,...,N_t\}$;
  Calculate $\mathbf{t}_k^i = \mathbf{H}_k \mathbf{p}_k^i$;
  Calculate $\left[\mathbf{q}_k^i\right]_m = \frac{1}{\sqrt{N_r^k}} e^{j\cdot\omega_m^i},\ m \in \left\{1,2,...,N_r^k\right\}$;
  Calculate $\tilde{\mathbf{H}}(k,k) = \left(\mathbf{q}_k^i\right)^H \mathbf{H}_k \mathbf{p}_k^i$;
**end for**
find $i^* = \underset{i\in\left\{1,2,...,N_r^k\right\}}{\arg\max}\ \tilde{\mathbf{H}}(k,k)$;
Output $\mathbf{f}_k^f = \mathbf{p}_k^{i^*}$, $\mathbf{w}_k = \mathbf{q}_k^{i^*}$.

**Difference from Antenna Selection:** We note the difference in our proposed approach from the antenna selection techniques. For each user $k$, each time the proposed method employs a specific row of $\mathbf{H}_k$ to obtain $\mathbf{p}_k^i$ and $\mathbf{q}_k^i$ accordingly, while the selected $\mathbf{f}_k^f$ that maximises the effective channel gain $\tilde{\mathbf{H}}(k,k)$ applies to all rows of $\mathbf{H}_k$ at the BS for the $k$-th user, which means that all the transmit antennas are active. At each user side, all the receive an-

tennas are active and employ the analog combiner $\mathbf{w}_k$ to obtain the received symbol for demodulation. Therefore, the proposed approach is not an antenna selection scheme but termed as a 'virtual path selection' approach. Furthermore, we note that the design of analog precoder and combiner is only dependent on the channel, and therefore no additional information exchange is required between the BS and users.

**Decoupled Design:** For the proposed joint design, both $\mathbf{F}_{RF}$ and $\mathbf{W}$ need to be calculated at the BS, while each user $k$ needs to calculate $\mathbf{F}_{RF}$ and $\mathbf{w}_k$. In this chapter, we further introduce a decoupled virtual path selection approach, where the BS only needs to calculate $\mathbf{F}_{RF}$, while the calculation of each $\mathbf{w}_k$ is solely conducted at the user side. For each user $k$, we first formulate $\mathbf{p}_k^i$ by (7.10), and we find $i^*$ that maximises $\mathbf{h}_k^i \mathbf{p}_k^i$, expressed as

$$\left(i^*, \mathbf{p}_k^{i^*}\right) = \underset{i \in \left\{1,2,\ldots,N_r^k\right\}}{\arg\max} \mathbf{h}_k^i \mathbf{p}_k^i, \tag{7.14}$$

which is equivalent to finding $\mathbf{h}_k^i$ with the largest norm. For each analog combiner, $\mathbf{w}_k = \mathbf{q}_k$ is then obtained by (7.12) based on $\mathbf{t}_k^* = \mathbf{H}_k \mathbf{p}_k^{i^*}$. The detailed algorithm is given in Algorithm 7.2.

**Algorithm 7.2** Decoupled Precoder and Combiner Design

**input** : $\mathbf{H}_k$
**output** : $\mathbf{f}_k^f$, $\mathbf{w}_k$
**for** $i = 1$ to $N_r^k$ **do**
  Calculate $\left[\mathbf{p}_k^i\right]_m = \frac{1}{\sqrt{N_t}} e^{j \cdot \left(-\varphi_m^i\right)}, m \in \{1,2,\ldots,N_t\}$;
**end for**
find $i^* = \underset{i \in \left\{1,2,\ldots,N_r^k\right\}}{\arg\max} \mathbf{h}_k^i \mathbf{p}_k^i$;
$\mathbf{f}_k^f = \mathbf{p}_k^{i^*}$, $\mathbf{t}_k^* = \mathbf{H}_k \mathbf{f}_k^{full}$;
Calculate $[\mathbf{w}_k]_m = \frac{1}{\sqrt{N_r^k}} e^{j \cdot \omega_m^*}$, $m \in \left\{1,2,\ldots,N_r^k\right\}$;
Output $\mathbf{f}_k^f$, $\mathbf{w}_k$.

**Low-Complexity Sub-optimal Design:** A low-complexity approach is further proposed, where a random path is selected to obtain the analog precoders and combiners. Without loss of generality, we employ $\mathbf{h}_k^1$ to obtain $\mathbf{f}_k$

and $\mathbf{w}_k$ for each user $k$. This approach removes the complexity of calculating the optimal $i^*$ and requires the least computational cost. The algorithm is given in Algorithm 7.3.

**Extension to Partially-Connected Structures:** When partially-connected structures are considered as in Fig. 2.9 (b), each $\mathbf{f}_k^p \in \mathbb{C}^{M_p \times 1}$. Similar to the design for the fully-connected structures, we design each $\mathbf{p}_k^i$ as the conjugate transpose of part of the $\mathbf{h}_k^i$, where for partially-connected structures, we have

$$\left[\mathbf{p}_k^i\right]_m = \frac{1}{\sqrt{N_t}} e^{j\cdot\left(-\varphi_m^i\right)},\ m \in \left\{(k-1)M_p+1,\cdots,kM_p\right\}. \tag{7.15}$$

The final analog precoder $\mathbf{f}_k^p$ and combiner $\mathbf{w}_k$ for partially-connected structures can be similarly obtained via the above three algorithms. It is observed in (7.15) that the analog precoder for partially-connected structures can only harvest part of the array gains, and therefore a performance gap in terms of the sum rate will be observed compared to the fully-connected structures, which will be numerically verified.

**Algorithm 7.3** Sub-optimal Precoder and Combiner Design

**input** : $\mathbf{h}_k^1$
**output** : $\mathbf{f}_k^f$, $\mathbf{w}_k$
Calculate $\left[\mathbf{p}_k^1\right]_m = \frac{1}{\sqrt{N_t}} e^{j\cdot\left(-\varphi_m^1\right)},\ m \in \{1,2,...,N_t\}$;
$\mathbf{f}_k^f$=$\mathbf{p}_k^1$;
Calculate $\mathbf{t}_k = \mathbf{H}_k\mathbf{f}_k^f$;
Calculate $[\mathbf{w}_k]_m = \frac{1}{\sqrt{N_r^k}} e^{j\cdot\omega_m},\ m \in \left\{1,2,...,N_r^k\right\}$;
Output $\mathbf{f}_k^f$ and $\mathbf{w}_k$.

**Baseband Design:** With $\mathbf{F}_{RF}$ and $\mathbf{W}$ obtained based on the introduced methods via virtual path selection, in the baseband we employ a ZF precoding to eliminate the multi-user interference in the off-diagonal of $\tilde{\mathbf{H}}$, expressed as

$$\mathbf{F}_{BB} = \frac{1}{f}\cdot\tilde{\mathbf{H}}^H\left(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\right)^{-1}, \tag{7.16}$$

where the scaling factor $f$ is to guarantee that the power constraint $\|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 = K$ is met.

### 7.1.4 Computational Cost Analysis

It can be observed that the complexity of the proposed technique is dominated by the calculations of $\tilde{\mathbf{H}}(k,k)$ and $\mathbf{t}_k^i = \mathbf{H}_k\mathbf{p}_k^i$. Note that $\mathbf{H}_k \in \mathbb{C}^{N_r^k \times N_t}$ and $\mathbf{p}_k^i \in \mathbb{C}^{N_t \times 1}$, and therefore the calculation of $\mathbf{q}_k^i$ in (7.12) involves $N_r^k N_t$ multiplications. The calculation of $\tilde{\mathbf{H}}(k,k)$ in Algorithm 7.1 involves $\left(N_r^k + N_r^k N_t\right)$ multiplications. This has to be calculated $N_r^k$ times for each user, both at the transmitter and the receiver, which leads to the total computational cost for the fully-connected system as:

$$\mathrm{C}_{\text{Joint}} = 2\sum_{k=1}^{K} N_r^k \left(N_r^k + 2N_r^k N_t\right). \tag{7.17}$$

The complexity of Algorithm 7.2, 7.3 can be similarly obtained as

$$\mathrm{C}_{\text{decoupled}} = 2\sum_{k=1}^{K} N_r^k \left(N_r^k N_t\right) + \sum_{k=1}^{K} N_r^k, \tag{7.18}$$

and

$$\mathrm{C}_{\text{sub-optimal}} = \sum_{k=1}^{K} N_r^k N_t. \tag{7.19}$$

For the hybrid precoding approaches based on partially-connected structures, the resulting computational costs can be efficiently obtained by substituting $N_t$ with $M$ in (7.17)-(7.19).

### 7.1.5 Numerical Results

In this section, the numerical results based on Monte Carlo simulations are presented. Without loss of generality, both ULAs and UPAs are employed with the antenna spacing $d = \frac{\lambda}{2}$, and the number of propagation paths $L_u = 6$. We assume the azimuth angles $\theta \in [0, 2\pi]$, and the elevation angles $\phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. For simplicity we assume $N_r^k = N_r^0$, $\forall k$. We compare the proposed schemes with existing codebook-based scheme [67], analog-only precoding,

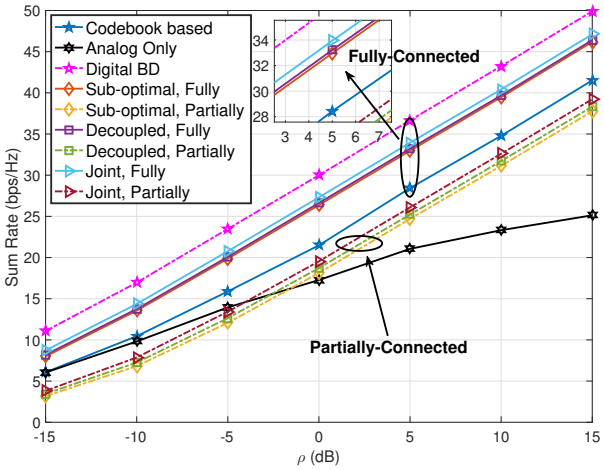



**Figure 7.1:** Sum rate v.s. transmit SNR, $N_t = 128$, $K = 4$, $N_r^0 = 16$, $L_u$=6, ULA

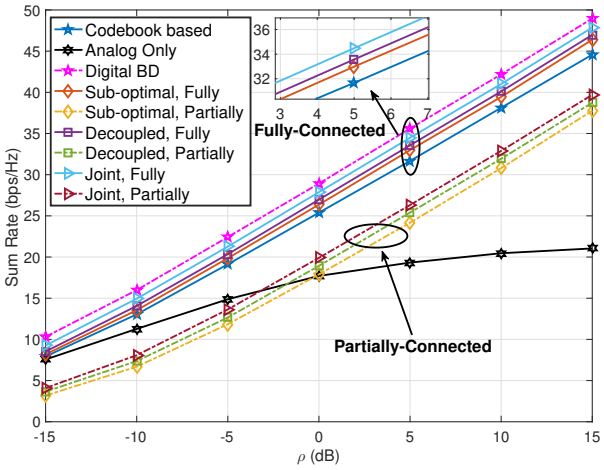



**Figure 7.2:** Sum rate v.s. transmit SNR, $N_t = 128$, $K = 4$, $N_r^0 = 16$, $L_u$=6, UPA

and FD BD precoding [43]. The transmit SNR in each figure is defined as $\rho = \frac{1}{K\sigma^2}$, where we have assumed a total transmit power $P = 1$.

Fig. 7.1 and Fig. 7.2 present the sum rate of different precoding methods with the increasing SNR for ULAs and UPAs respectively, where the the quantisation of the azimuth and elevation angle is 4 bits and 3 bits respectively for the codebook-based approach in [67]. For fully-connected structures, it is observed that all of the proposed methods outperform the existing codebook-based scheme and analog precoding, and can approach the performance of FD BD precoding. Specifically, in the case of UPAs we ob-

serve only a SNR loss of less than 2dB for our proposed technique. A similar trend can be observed for the partially-connected structures, where the joint design achieves the best sum rate performance, followed by the proposed decoupled method. For partially-connected structures, there is a loss in sum rates as only part of the array gain can be obtained.

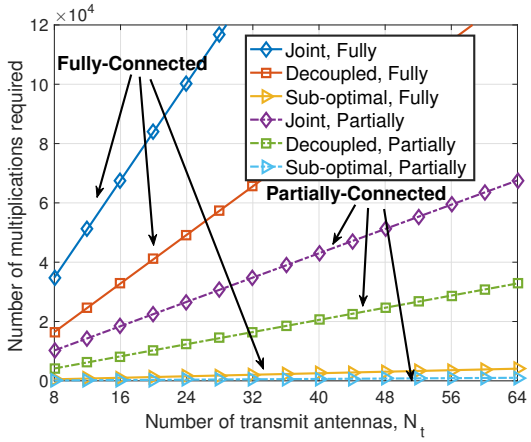



**Figure 7.3:** Number of multiplications required v.s. number of transmit antennas, $K = 4$, $N_r^0 = 16$, $L_u$=6, UPA

Fig. 7.3 compares the computational cost for each proposed algorithm in terms of the multiplication number required. In general, it is observed that fewer multiplications are needed for partially-connected structures, as $N_t$ is substituted by $M_p$. For both fully-connected and partially-connected structures, it is observed that the joint design requires the most number of multiplications, while as expected the computational cost for the low-complexity sub-optimal approach is negligible, compared to the joint and decoupled design.

## 7.2 Hybrid Precoding for Energy-Efficient SWIPT

In this section, we extend the study on hybrid precoding to MIMO SWIPT systems for an energy-efficient transmission.

### 7.2.1 Introduction

With the increasing traffic demand and number of user equipments (UEs) in the wireless environment, the power consumption of both the BSs and UEs in

the wireless communication systems has increased dramatically [145, 146]. However, most UEs only have limited power supplies (batteries) currently, which has become a bottleneck when the power consumption becomes increasingly high, and the slow development speed of battery techniques cannot satisfy the current energy requirement [147]. Towards this direction, energy harvesting (EH) techniques have been proposed to exploit the natural energy such as solar, tide and wind to prolong the battery life of UEs [148]. However, such techniques usually depend on the environmental conditions and the natural energy may not always be available, especially for the indoor environments.

Recent advances have shown that electromagnetic (EM) radiation can be exploited as a potential energy source, based on the fact that the energy included in EM waves can be converted to direct current (DC) voltage with rectenna circuits [149, 150]. A step further has been obtained by the wireless power transfer technique, which has been extensively studied for wireless sensor networks [151, 152, 153, 154]. Similarly for wireless communications, the energy harvesting techniques and wireless power transfer enable the UEs to harvest energy from the EM waves in the communication links, and therefore have become particularly appealing [155, 156, 157, 158]. For wireless communication systems, the RF signals carry both the information and energy at the same time, and there exists a fundamental tradeoff between information decoding and energy harvesting, which has been studied in [157] for flat fading and [158] for selective fading, respectively. Nevertheless, an ideal receiver that can simultaneously decode information and harvest energy with the same received signal is assumed in [157, 158], which is not applicable currently. Therefore, a more practical approach termed as SWIPT is considered in [159], where three different types of receivers are proposed, i.e., separate, time-switching and power-splitting. MIMO SWIPT techniques have recently attracted a lot of research attention [160, 161, 162, 163, 164], where in [161, 162] the SWIPT techniques for cellular networks

are considered, and in [163] a robust precoder for MIMO SWIPT systems with stochastic Rician fading is proposed. [164] employs the ZF precoding method for SWIPT systems, and shows that the energy obtained by the EH receivers can be increased at the cost of a SINR loss of the information decoders (IDs). The downlink precoding scheme that maximises the received energy for EH receivers while guaranteeing the SINR of the IDs is considered in [165]. A data-aided downlink precoding that exploits the CI is proposed in [166] to further improve the performance of MIMO SWIPT systems. In [167, 168, 169, 170], the joint information and energy precoding approaches for MIMO interference channels have been investigated, while the study of MIMO SWIPT techniques has also been combined with physical layer security in [171, 172, 173, 174] by considering the broadcast nature of the wireless communications.

In addition to the SWIPT techniques, energy-efficient transmission is another way to manage the increasing power consumption of the wireless industry. The above existing designs for SWIPT with MIMO systems all assume a FD precoder, which requires a dedicated pair of DACs and a RF chain for each antenna element. Such FD design for MIMO systems will lead to a high power consumption at the BS. Even with moderate numbers of antennas, the power consumption of the RF chains is dominant, and the hybrid AD precoding techniques that allow a reduction in the RF chains are desirable. While the hybrid AD precoding has been widely studied for large-scale MIMO systems, what has been neglected in the existing literature is that the hybrid structure is a promising candidate for energy-efficient transmission, which meets the requirement for the future wireless communication systems. Moreover, the reduction in the hardware complexity and power consumption directly applies to small-scale MIMO systems. Indeed, small APs for the future IoTs or small BSs (for example femtocells or picocells that are widely deployed for heterogeneous networks) usually have limited power supply, which can benefit from the hybrid structures. Interestingly, such

power-efficient approaches by hybrid precoding have yet to be explored for SWIPT.

Accordingly, in this chapter we investigate the SWIPT techniques for small-scale MIMO systems with limited RF chains at the APs, where energy-efficient hybrid AD precoding is considered. We study the scenario where the BS serves one ID and several EH receivers simultaneously, and we focus on the minimisation of the total transmit power at the BS, while meeting the SINR requirement of the ID and the harvested energy requirement of each EH receiver. Firstly, we mathematically analyse the FD precoding problem with Lagrangian and Karush-Kuhn-Tucker (KKT) conditions, and analytically show that the optimality is achieved by employing the information precoder only. Specifically, for the case where there is only one EH receiver, we obtain the closed-form expressions of the optimal precoding vectors. The above study mathematically proves that the optimality can be achieved by employing the information precoder only. We proceed to consider the hybrid case, where we firstly propose a low-complexity hybrid approach as a performance benchmark. For the low-complexity hybrid scheme, the analog precoders are obtained based on the SVD of the channel, and the low-dimensional digital precoder is subsequently obtained based on the effective analog channel. To improve upon the above hybrid approach, an iterative hybrid method is further introduced, where in each iteration we design the analog precoder by minimising the Euclidean distance between the FD precoder and the hybrid precoder. Based on our analyses on the FD case, the optimal analog precoder can be efficiently solved via a geometrical interpretation. For both hybrid precoding schemes, the extension to partially-connected structures at the BS is also introduced. The numerical results show that the proposed iterative method achieves a near-optimal performance for fully-connected structures, while the performance gains over the hybrid case based on SVD are more significant for partially-connected structures. It is important to observe that the hybrid structures consume

much less total power at the BS to achieve the same performance as the FD structure, which verifies that the hybrid structures are more favourable for the future energy-efficient transmission.

For reasons of clarity, we summarise the contributions of this paper as:

- We analyse the FD MIMO SWIPT power minimisation problem with Lagrangian and KKT approaches, where we show analytically that the optimality is achieved by employing the information precoder only. For the special case with only one EH receiver in the system, we derive the closed-form expressions of the optimal FD precoders.
- Based on the above observation, we proceed to consider a small-scale MIMO SWIPT system with limited RF chains at the BS, and propose a low-complexity hybrid precoding method for the power minimisation problem, where the analog precoders are obtained based on SVD.
- We further propose an iterative hybrid scheme, where the analog precoder is designed by minimising the Euclidean distance between the hybrid precoder and the FD precoder. Within each iteration, the optimal solution of the analog precoder can be efficiently obtained via a geometrical approach.

### 7.2.2 System Model and FD SWIPT

We consider a downlink MIMO system as shown in Fig. 7.4, where a BS with $N_t$ antennas and $N_{RF}^t$ RF chains serves one single-antenna ID and $K$ single-antenna EH receivers simultaneously. We assume that perfect CSI is available at the BS throughout this chapter. A spatially-uncorrelated flat-fading Rayleigh MIMO channel is assumed, and we denote $\mathbf{h}_I \in \mathbb{C}^{1\times N_t}$ and $\mathbf{h}_E^k \in \mathbb{C}^{1\times N_t}$ as the channel from the BS to the ID and EH receiver $k$, respectively. Each entry of $\mathbf{h}_I$ and $\mathbf{h}_E^k$ is modelled as

$$\begin{aligned} [\mathbf{h}_I]_m &= \sqrt{\alpha_0 D_I^{-\beta} C_I} \cdot [\mathbf{g}_I]_m, \; m \in \mathscr{N}, \\ \left[\mathbf{h}_E^k\right]_m &= \sqrt{\alpha_0 \left(D_E^k\right)^{-\beta} C_E^k} \cdot \left[\mathbf{g}_E^k\right]_m, \; m \in \mathscr{N}, \end{aligned} \tag{7.20}$$

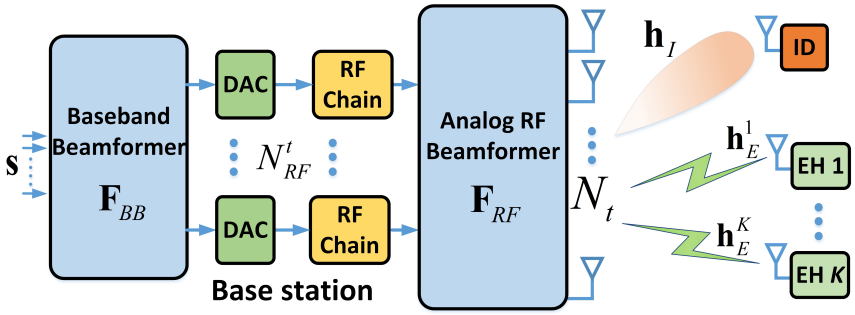



**Figure 7.4:** Block diagram of the hybrid AD precoding for SWIPT

which is valid for far-field communication where the distance between the receivers and the BS is on the order of tens of meters [175], and we denote $\mathcal{N} = \{1, 2, ..., N_t\}$. In (7.20), $\alpha_0$ is a constant determined by the wireless propagation environment, $D_I$ is the distance between the BS and the ID, $\beta$ represents the pathloss coefficient, and $C_I$ denotes the shadow fading. Each element in $\mathbf{g}_I$ is independent and follows the standard complex Gaussian distribution, which forms the Rayleigh component of the channel. The denotation is similar for each $\mathbf{h}_E^k$ of the EH receiver $k$.

When a conventional FD precoding is considered, $N^t_{RF} = N_t$ and we denote the corresponding FD precoding matrix as $\mathbf{W} = \left[\mathbf{w}_I, \mathbf{w}_E^1, ..., \mathbf{w}_E^K\right]$. Accordingly, we can express the received symbol at the ID as

$$r_I = \mathbf{h}_I \mathbf{w}_I s_I + \mathbf{h}_I \sum_{i=1}^{K} \mathbf{w}_E^i s_E^i + n_I, \tag{7.21}$$

where $s_I$ and each $s_E^i$ denote the data symbol. $n_I$ represents the additive Gaussian noise with zero mean and variance $\sigma^2$. Following the existing literature [165, 167, 170], we express the received SINR for the ID as

$$\gamma_I = \frac{\left|\mathbf{h}_I \mathbf{w}_I\right|^2}{\sum\limits_{k=1}^{K} \left|\mathbf{h}_I \mathbf{w}_E^k\right|^2 + \sigma^2}. \tag{7.22}$$

For the EH receivers, we assume the noise power at each receiver is the same as that of ID. For simplicity, we assume a linear energy harvesting model

[165, 170, 176], based on which we can obtain the harvested energy for the $k$-th EH receiver as

$$E_k = \eta \left( \left| \mathbf{h}_E^k \mathbf{w}_I \right|^2 + \sum_{i=1}^{K} \left| \mathbf{h}_E^k \mathbf{w}_E^i \right|^2 + \sigma^2 \right), \tag{7.23}$$

where $\eta$ is a constant that represents the efficiency of converting the received radio signals into electrical energy. In (7.23), $0 < \eta < 1$ and for simplicity we have assumed an identical energy transformation efficiency $\eta$ for each user.

We consider the optimisation problem where the transmit power is minimised while meeting the SINR requirement of the ID and the harvested energy requirement of each EH receiver, which can be constructed as [166]

$$\begin{aligned} \mathscr{P}_{7.1}: \ & \min_{\mathbf{w}_I, \mathbf{w}_E^i} \ p \\ s.t. \quad & p \geq \|\mathbf{w}_I\|^2 + \sum_{i=1}^{K} \left\|\mathbf{w}_E^i\right\|^2 \\ & \gamma_I \geq \gamma_0, \ E_k \geq E_0, \ \forall k \in \mathscr{K} \end{aligned} \tag{7.24}$$

where $\mathscr{K} = \{1, 2, ..., K\}$, and $\gamma_0$ is the SINR requirement of the ID. For simplicity we have assumed an identical harvested energy requirement for each EH receiver, which is denoted as $E_0$.

### 7.2.3 Analyses on the FD SWIPT

To introduce the rationale behind the proposed hybrid iterative precoding in Chapter 7.2.5, we firstly perform mathematical analyses on the FD precoding problem with the Lagrangian and KKT approach, where we show that the optimality of $\mathscr{P}_{7.1}$ is achieved by employing the information precoder only. While the KKT conditions are only necessary conditions for non-convex optimisation problems, for the considered problems we can verify that the obtained solutions are also sufficient. Similar conclusions can be drawn for the hybrid case, and these two observations motivate the design of the proposed iterative algorithm in Chapter 7.2.5.

#### 7.2.3.1 The Case of Only One EH Receiver

Firstly, we consider a special case where there is one ID and one EH receiver in the system, and in this case $\mathscr{P}_{7.1}$ can be simplified as

$$\begin{aligned} \mathscr{P}_{7.2}: \ & \min_{\mathbf{w}_I, \mathbf{w}_E^i} \ \mathbf{w}_I^H \mathbf{w}_I + \mathbf{w}_E^H \mathbf{w}_E \\ s.t. \quad & \mathbf{h}_I \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_I^H + \sigma^2 - \frac{1}{\gamma_0} \mathbf{h}_I \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_I^H \leq 0 \\ & \frac{E_0}{\eta} - \sigma^2 - \mathbf{h}_E \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_E^H - \mathbf{h}_E \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_E^H \leq 0 \end{aligned} \tag{7.25}$$

We express the Lagrangian of $\mathscr{P}_{7.2}$ as [124]

$$\begin{aligned} \mathscr{L}(\mathbf{w}_I, \mathbf{w}_E, \lambda_I, \lambda_E) = \mathbf{w}_I^H \mathbf{w}_I + \mathbf{w}_E^H \mathbf{w}_E \lambda_I \left( \mathbf{h}_I \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_I^H + \sigma^2 - \frac{1}{\gamma_0} \mathbf{h}_I \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_I^H \right) \\ + \lambda_E \left( \frac{E_0}{\eta} - \sigma^2 - \mathbf{h}_E \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_E^H - \mathbf{h}_E \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_E^H \right), \end{aligned} \tag{7.26}$$

where $\lambda_I$ and $\lambda_E$ denote the dual variables with respect to the SINR and harvested energy constraint, respectively. Accordingly, the KKT conditions for optimality are obtained as

$$\frac{\partial \mathscr{L}}{\partial \mathbf{w}_I} = \mathbf{w}_I^H - \frac{\lambda_I}{\gamma_0} \mathbf{w}_I^H \mathbf{h}_I^H \mathbf{h}_I - \lambda_E \mathbf{w}_I^H \mathbf{h}_E^H \mathbf{h}_E = \mathbf{0} \tag{7.27a}$$

$$\frac{\partial \mathscr{L}}{\partial \mathbf{w}_E} = \mathbf{w}_E^H + \lambda_I \mathbf{w}_E^H \mathbf{h}_I^H \mathbf{h}_I - \lambda_E \mathbf{w}_E^H \mathbf{h}_E^H \mathbf{h}_E = \mathbf{0} \tag{7.27b}$$

$$\lambda_I \left( \mathbf{h}_I \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_I^H + \sigma^2 - \frac{1}{\gamma_0} \mathbf{h}_I \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_I^H \right) = 0 \tag{7.27c}$$

$$\lambda_E \left( \frac{E_0}{\eta} - \sigma^2 - \mathbf{h}_E \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_E^H - \mathbf{h}_E \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_E^H \right) = 0 \tag{7.27d}$$

In the following, dependent on whether the SINR and energy constraint are active or not, we discuss the optimality condition of the optimisation problem, and obtain the closed-form expressions of the optimal precoding vectors and the corresponding transmit power $P_{TX}$ required.

**Only SINR Constraint is Active:** When only the SINR constraint is strictly met, this corresponds to the case where the SINR target is high and

more demanding compared to the harvested energy requirement, which leads to the following proposition.

**Proposition 1:** When only the SINR constraint is active, the optimal precoding vectors are given by

$$\mathbf{w}_I^* = \frac{\sqrt{\gamma_0 \sigma^2}}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)} \cdot \mathbf{h}_I^H, \ \mathbf{w}_E^* = \mathbf{0}, \tag{7.28}$$

the energy target $E_0$ of the EH receiver should satisfy

$$E_0 \leq \frac{\gamma_0 \eta \sigma^2 \left(\mathbf{h}_I \mathbf{h}_E^H \mathbf{h}_E \mathbf{h}_I^H\right)}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)^2} + \eta \sigma^2, \tag{7.29}$$

and the corresponding transmit power required is

$$P_{TX}^* = \frac{\gamma_0 \sigma^2}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)}. \tag{7.30}$$

**Proof:** In this case, the harvested energy by the information precoder already meets the energy requirement for the EH receiver, which leads to

$$\eta \left(\mathbf{h}_E \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_E^H + \sigma^2\right) > E_0, \tag{7.31}$$

and $\mathbf{w}_E^* = \mathbf{0}$ because in this case the presence of the energy precoder will only degrade the received SINR performance of the ID. Based on the complementary slackness condition in (7.27d), we can further obtain $\lambda_I > 0$ and $\lambda_E = 0$. Based on the above, the KKT conditions can be simplified into

$$\mathbf{w}_I^H - \frac{\lambda_I}{\gamma_0} \mathbf{w}_I^H \mathbf{h}_I^H \mathbf{h}_I = \mathbf{0}, \ \ \sigma^2 - \frac{1}{\gamma_0} \mathbf{h}_I \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_I^H = 0. \tag{7.32}$$

We note that the optimal $\mathbf{w}_I$ is not unique and is invariant to a phase rotation, and therefore in (7.32) we can assume there exists an optimal $\mathbf{w}_I$ such that $\Im\left\{\mathbf{h}_I \mathbf{w}_I\right\} = \Im\left\{\mathbf{w}_I^H \mathbf{h}_I^H\right\} = 0$, which further leads to

$$\mathbf{h}_I \mathbf{w}_I = \mathbf{w}_I^H \mathbf{h}_I^H = \sqrt{\gamma_0 \sigma^2}. \tag{7.33}$$

By substituting (7.33) into (7.32), we can further obtain $\lambda_I = \frac{\gamma_0}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)}$, which leads to the final expression of the optimal $\mathbf{w}_I$, given by

$$\mathbf{w}_I^* = \frac{\sqrt{\gamma_0 \sigma^2}}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)} \cdot \mathbf{h}_I^H. \tag{7.34}$$

Accordingly, we can obtain the transmit power as

$$P_{TX}^* = \left(\mathbf{w}_I^*\right)^H \mathbf{w}_I^* = \frac{\gamma_0 \sigma^2}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)}, \tag{7.35}$$

which is only relevant to the SINR target of the ID and not relevant to $E_0$. To fall within this category, the harvested energy requirement should satisfy

$$\frac{E_0}{\eta} - \sigma^2 \leq \mathbf{h}_E \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_E^H \;\Rightarrow\; E_0 \leq \frac{\gamma_0 \eta \sigma^2 \left(\mathbf{h}_I \mathbf{h}_E^H \mathbf{h}_E \mathbf{h}_I^H\right)}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)^2} + \eta \sigma^2 \triangleq E_{th}^1. \tag{7.36}$$

**Only Energy Constraint is Active:** When only the energy constraint is active, this corresponds to the case where the harvested energy requirement is more demanding, which leads to the following proposition.

**Proposition 2:** When only the energy constraint is active, the optimal precoding vectors are given by

$$\mathbf{w}_I^* = \frac{\sqrt{\frac{E_0 - \eta \sigma^2}{\eta\left(1+c^2\right)}}}{\left(\mathbf{h}_E \mathbf{h}_E^H\right)} \cdot \mathbf{h}_E^H, \;\; \mathbf{w}_E^* = c \cdot \mathbf{w}_I^*, \tag{7.37}$$

where $c \geq 0$ is real and satisfies (7.46) which is shown in the proof that follows. The harvested energy requirement $E_0$ should satisfy

$$E_0 \geq \frac{\gamma_0 \eta \sigma^2 \left(\mathbf{h}_E \mathbf{h}_E^H\right)^2}{\left(\mathbf{h}_I \mathbf{h}_E^H \mathbf{h}_E \mathbf{h}_I^H\right)} + \eta \sigma^2, \tag{7.38}$$

and the corresponding transmit power required is

$$P_{TX}^* = \frac{E_0 - \eta \sigma^2}{\eta \left(\mathbf{h}_E \mathbf{h}_E^H\right)}. \tag{7.39}$$

**Proof:** In this case, the SINR requirement for the ID is over-satisfied, and we obtain

$$\frac{1}{\gamma_0}\mathbf{h}_I\mathbf{w}_I\mathbf{w}_I^H\mathbf{h}_I^H > \mathbf{h}_I\mathbf{w}_E\mathbf{w}_E^H\mathbf{h}_I^H + \sigma^2. \tag{7.40}$$

Based on the complementary slackness condition in (7.27c), we further obtain $\lambda_I = 0$ and $\lambda_E > 0$, and the stationarity conditions in (7.27a) and (7.27b) can be transformed into

$$\mathbf{w}_I^H - \lambda_E\mathbf{w}_I^H\mathbf{h}_E^H\mathbf{h}_E = \mathbf{w}_I^H\left(\mathbf{I} - \lambda_E\mathbf{h}_E^H\mathbf{h}_E\right) = \mathbf{0} \tag{7.41a}$$

$$\mathbf{w}_E^H - \lambda_E\mathbf{w}_E^H\mathbf{h}_E^H\mathbf{h}_E = \mathbf{w}_E^H\left(\mathbf{I} - \lambda_E\mathbf{h}_E^H\mathbf{h}_E\right) = \mathbf{0} \tag{7.41b}$$

It is observed that the optimal $\mathbf{w}_I$ and $\mathbf{w}_E$ are parallel, and without loss of generality we can assume

$$\mathbf{w}_E = c\cdot\mathbf{w}_I, \tag{7.42}$$

where $c \geq 0$. Furthermore, as the energy constraint is active, we substitute (7.42) into (7.27d), which yields

$$\left(1+c^2\right)\mathbf{h}_E\mathbf{w}_I\mathbf{w}_I^H\mathbf{h}_E^H = \frac{E_0}{\eta} - \sigma^2 \;\Rightarrow\; \mathbf{h}_E\mathbf{w}_I = \mathbf{w}_I^H\mathbf{h}_E^H = \sqrt{\frac{E_0 - \eta\sigma^2}{\eta\left(1+c^2\right)}}, \tag{7.43}$$

By substituting (7.43) into (7.41a), the optimal information precoder and energy precoder can be obtained as

$$\mathbf{w}_I^* = \frac{\sqrt{\frac{E_0-\eta\sigma^2}{\eta\left(1+c^2\right)}}}{\left(\mathbf{h}_E\mathbf{h}_E^H\right)}\cdot\mathbf{h}_E^H,\;\; \mathbf{w}_E^* = \frac{c\sqrt{\frac{E_0-\eta\sigma^2}{\eta\left(1+c^2\right)}}}{\left(\mathbf{h}_E\mathbf{h}_E^H\right)}\cdot\mathbf{h}_E^H. \tag{7.44}$$

Incorporating the above expressions of the optimal precoding vectors into (7.31), we can obtain

$$\frac{1}{\gamma_0}\frac{\frac{E_0-\eta\sigma^2}{\eta\left(1+c^2\right)}}{\left(\mathbf{h}_E\mathbf{h}_E^H\right)^2}\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right) > c^2\frac{\frac{E_0-\eta\sigma^2}{\eta\left(1+c^2\right)}}{\left(\mathbf{h}_E\mathbf{h}_E^H\right)^2}\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right) + \sigma^2. \tag{7.45}$$

With some further transformations, we obtain that $c$ should satisfy

$$\begin{aligned} &\frac{1}{\gamma_0} > c^2 + \frac{\left(1+c^2\right)\eta\sigma^2\left(\mathbf{h}_E\mathbf{h}_E^H\right)^2}{\left(E_0-\eta\sigma^2\right)\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right)} \\ \Rightarrow\ &c^2 < \frac{1}{\gamma_0}\cdot\frac{\left(E_0-\eta\sigma^2\right)\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right)-\gamma_0\eta\sigma^2\left(\mathbf{h}_E\mathbf{h}_E^H\right)^2}{\left(E_0-\eta\sigma^2\right)\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right)+\eta\sigma^2\left(\mathbf{h}_E\mathbf{h}_E^H\right)^2}. \end{aligned} \tag{7.46}$$

Since $c^2 \geq 0$, based on (7.46) we can obtain the requirement for $E_0$, which is given by

$$E_0 \geq \frac{\gamma_0\eta\sigma^2\left(\mathbf{h}_E\mathbf{h}_E^H\right)^2}{\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right)}+\eta\sigma^2 \triangleq E_{th}^2, \tag{7.47}$$

and we further note that $E_{th}^2$ obtained in (7.47) is guaranteed to be larger than $E_{th}^1$ in (7.36) based on the inner-product property, where we have

$$\mathbf{h}_I\mathbf{h}_I^H\mathbf{h}_E\mathbf{h}_E^H \geq \mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H. \tag{7.48}$$

In (7.48), the equality holds only when $\mathbf{h}_I$ and $\mathbf{h}_E$ are parallel. Furthermore, we can obtain the total power consumption as

$$P_{TX}^* = \left(1+c^2\right)\left(\mathbf{w}_I^*\right)^H\mathbf{w}_I^* = \frac{E_0-\eta\sigma^2}{\eta\left(\mathbf{h}_E\mathbf{h}_E^H\right)}, \tag{7.49}$$

which is independent of $c$. This can also be observed from the optimisation problem itself, as the EH receiver harvests the energy from both the information precoder and the energy precoder. As long as the value of $c$ satisfies (7.46) (this guarantees that the SINR target of the ID is met), how the power is distributed between $\mathbf{w}_I$ and $\mathbf{w}_E$ according to (7.44) will not have an impact on the total amount of energy harvested by the EH receiver. We note that by choosing $c = 0$, the optimality can be achieved by employing the information precoder only.

**Both Constraints are Active:** When both the SINR constraint and energy constraint are active, we first note that if $\lambda_E = 0$, the optimal solution will be the same as (7.28), which corresponds to the extreme point before

which the energy constraint is not active, and the optimal solution is to employ information precoder only. If $\lambda_E = 0$, the optimal solution will be (7.37). Therefore in the following, we focus on the case where $\lambda_I > 0$ and $\lambda_E > 0$, which leads to the following proposition.

**Proposition 3:** When both the SINR constraint and energy constraint are active, the optimal precoding vectors can be expressed as

$$\mathbf{w}_I^* = \alpha \cdot \mathbf{h}_I^H + \beta \cdot \mathbf{h}_\perp^H, \ \ \mathbf{w}_E^* = \mathbf{0}, \tag{7.50}$$

where $\mathbf{h}_\perp = \mathbf{h}_E - \frac{\mathbf{h}_E \mathbf{h}_I^H \mathbf{h}_I}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)}$ is orthogonal to $\mathbf{h}_I$, and $\alpha, \beta$ are the weighting factors. The energy requirement for the EH receiver should satisfy

$$\frac{\gamma_0 \eta \sigma^2 \left(\mathbf{h}_I \mathbf{h}_E^H \mathbf{h}_E \mathbf{h}_I^H\right)}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)^2} + \eta \sigma^2 < E_0 < \frac{\gamma_0 \eta \sigma^2 \left(\mathbf{h}_E \mathbf{h}_E^H\right)^2}{\left(\mathbf{h}_I \mathbf{h}_E^H \mathbf{h}_E \mathbf{h}_I^H\right)} + \eta \sigma^2, \tag{7.51}$$

and the corresponding transmit power required is

$$P_{TX}^* = \lambda_I \sigma^2 + \lambda_E \left( \frac{E_0}{\eta} - \sigma^2 \right). \tag{7.52}$$

**Proof:** We firstly derive the expression of the total transmit power, based on which we prove that $\mathbf{w}_E^* = \mathbf{0}$ by contradiction, and subsequently we obtain the expression of the optimal information precoder $\mathbf{w}_I^*$. To be specific, we multiply $\mathbf{w}_I$ and $\mathbf{w}_E$ to the right-hand side of (7.27a) and (7.27b) respectively, and we can further obtain

$$\mathbf{w}_I^H \mathbf{w}_I = \frac{\lambda_I}{\gamma_0} \mathbf{h}_I \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_I^H + \lambda_E \mathbf{h}_E \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_E^H \tag{7.53a}$$

$$\mathbf{w}_E^H \mathbf{w}_E = -\lambda_I \mathbf{h}_I \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_I^H + \lambda_E \mathbf{h}_E \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_E^H \tag{7.53b}$$

The sum of (7.53a) and (7.53b) yields

$$\begin{aligned} \mathbf{w}_I^H \mathbf{w}_I + \mathbf{w}_E^H \mathbf{w}_E = \ & \lambda_I \left( \frac{1}{\gamma_0} \mathbf{h}_I \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_I^H - \mathbf{h}_I \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_E^H \right) \\ & + \lambda_E \left( \mathbf{h}_E \mathbf{w}_I \mathbf{w}_I^H \mathbf{h}_E^H + \mathbf{h}_E \mathbf{w}_E \mathbf{w}_E^H \mathbf{h}_E^H \right). \end{aligned} \tag{7.54}$$

It is observed in (7.54) that the left-hand side is the total transmit power. Since both the SINR constraint and the energy constraint are active, the right-hand side can be further simplified and the total transmit power is obtained as

$$P_{TX}^{*} = \lambda_I \sigma^2 + \lambda_E \left( \frac{E_0}{\eta} - \sigma^2 \right), \tag{7.55}$$

which means that the optimal transmit power is only related to the dual variables.

Based on (7.55), we proceed to prove that $\mathbf{w}_E^{*} = \mathbf{0}$ by contradiction. We firstly assume one case, where the optimal solution is an information precoder $\mathbf{w}_I^0$ only that satisfies both of the constraints, and the corresponding power consumption is obtained as

$$P_{TX}^{0} = \lambda_I^0 \sigma^2 + \lambda_E^0 \left( \frac{E_0}{\eta} - \sigma^2 \right). \tag{7.56}$$

In addition, we assume another case, where we need an information precoder $\mathbf{w}_I$ and an energy precoder $\mathbf{w}_E$ to satisfy both of the constraints, with the same total transmit power. The total power consumption in this case is expressed as

$$P_{TX} = \lambda_I \sigma^2 + \lambda_E \left( \frac{E_0}{\eta} - \sigma^2 \right). \tag{7.57}$$

As can be observed, $P_{TX} = P_{TX}^0$ is equivalent to $\lambda_I = \lambda_I^0$ and $\lambda_E = \lambda_E^0$. Subsequently, we express the stationarity condition for $\mathbf{w}_I^0$ and $\mathbf{w}_I$ as

$$\left(\mathbf{w}_I^0\right)^H \left( \mathbf{I} - \frac{\lambda_I^0}{\gamma_0} \mathbf{h}_I^H \mathbf{h}_I - \lambda_E^0 \mathbf{h}_E^H \mathbf{h}_E \right) = \mathbf{0}, \tag{7.58a}$$

$$\mathbf{w}_I^H \left( \mathbf{I} - \frac{\lambda_I}{\gamma_0} \mathbf{h}_I^H \mathbf{h}_I - \lambda_E \mathbf{h}_E^H \mathbf{h}_E \right) = \mathbf{0}. \tag{7.58b}$$

With $\lambda_I = \lambda_I^0$ and $\lambda_E = \lambda_E^0$, it is observed that $\mathbf{w}_I^0$ and $\mathbf{w}_I$ are parallel, and without loss of generality we assume

$$\mathbf{w}_I = a \cdot \mathbf{w}_I^0, \tag{7.59}$$

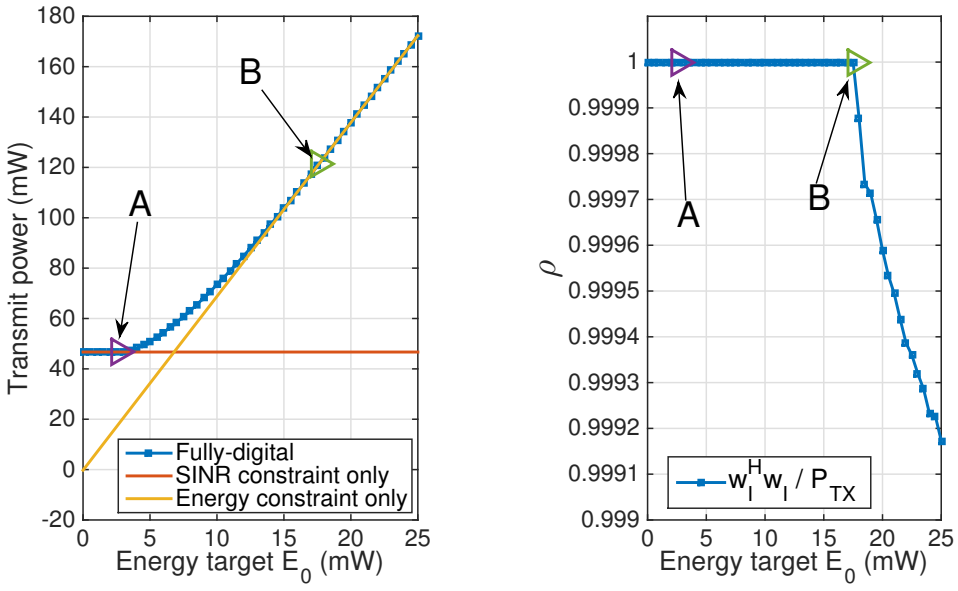
where $a$ is real and $a \neq 0$. As the SINR constraint is active for both cases, we can obtain

$$\begin{aligned} \frac{1}{\gamma_0}\mathbf{h}_I\mathbf{w}_I^0\left(\mathbf{w}_I^0\right)^H\mathbf{h}_I^H &= \sigma^2, \\ \frac{a^2}{\gamma_0}\mathbf{h}_I\mathbf{w}_I^0\left(\mathbf{w}_I^0\right)^H\mathbf{h}_I^H &= \mathbf{h}_I\mathbf{w}_E\mathbf{w}_E^H\mathbf{h}_I^H + \sigma^2. \end{aligned} \tag{7.60}$$

(7.60) can be further transformed into

$$\frac{\left(a^2-1\right)}{\gamma_0}\mathbf{h}_I\mathbf{w}_I^0\left(\mathbf{w}_I^0\right)^H\mathbf{h}_I^H = \mathbf{h}_I\mathbf{w}_E\mathbf{w}_E^H\mathbf{h}_I^H. \tag{7.61}$$

Since $\mathbf{w}_E \neq 0$, we have $a^2 > 1$. Following a similar step to (7.60), (7.61) for the energy constraint, we obtain $a^2 < 1$ such that the energy constraint is met, which causes contradiction. Therefore, the optimal case is to employ the information precoder only.




**(a)** Transmit power (mW) **(b)** Value of $\rho$

**Figure 7.5:** $N_t = 12$, one ID, $K = 1$ EH receiver, $\gamma_0 = 16$dB, $D_I$=10m, $D_E$=5m, $\eta$=0.35

To validate the above analyses, in Fig. 7.5 (a) and Fig. 7.5 (b) we show the transmit power and the power ratio $\rho$ with respect to the harvested energy requirement $E_0$ respectively, where $\rho$ represents the percentage that the

information precoder accounts for the total transmit power and is defined as

$$\rho = \frac{\mathbf{w}_I^H \mathbf{w}_I}{P_{TX}}. \tag{7.62}$$

In Fig. 7.5, the point 'A' denotes the extreme point before which the harvested energy constraint is not active, and $E_0^A = E_{th}^1$. The point 'B' is the extreme point after which the energy precoder can be introduced, and $E_0^B = E_{th}^2$. When $E_0 \leq E_{th}^1$, there is only an information precoder and only the SINR constraint is active, in which case the required transmit power remains constant; When $E_{th}^1 < E_0 < E_{th}^2$, both of the constraints are active; When $E_0 \geq E_{th}^2$, the energy precoder can be introduced, as validated in Fig. 7.5 (b), and the required transmit power is linearly increasing with the increasing harvested energy requirement $E_0$, given by (7.49).

When both of the SINR and harvested energy constraints are active, it is difficult to compute the exact closed-form expression of $\mathbf{w}_I^*$. Nevertheless, based on the observation that both of the SINR and energy constraints are active with the increasing $E_0$ when $E_0 \in \left[E_{th}^1, E_{th}^2\right]$, we can obtain that the optimal precoding vector $\mathbf{w}_I$ is in the form of [177]

$$\mathbf{w}_I = \alpha \cdot \mathbf{h}_I^H + \beta \cdot \mathbf{h}_\perp^H, \tag{7.63}$$

where $\mathbf{h}_\perp$ is orthogonal to $\mathbf{h}_I$ and can be expressed as $\mathbf{h}_\perp = \mathbf{h}_E - \frac{\mathbf{h}_E \mathbf{h}_I^H \mathbf{h}_I}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)}$. In (7.63), $\alpha$ can be chosen as $\alpha = \frac{\sqrt{\gamma_0 \sigma^2}}{\left(\mathbf{h}_I \mathbf{h}_I^H\right)}$. This structure ensures that the SINR constraint is met, and the value of the complex weighting factor $\beta$ dependent on $E_0$ can be obtained with the active energy harvesting constraint.

We summarise the optimal information precoder, energy precoder and the corresponding transmit power for the case of one ID and one EH receiver in Table 7.1.

| Condition of $E_0$ | $\mathbf{w}_I^*$ | $\mathbf{w}_E^*$ | $P_{TX}^*$ |
|---|---|---|---|
| $E_0 \leq \frac{\gamma_0\eta\sigma^2\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right)}{\left(\mathbf{h}_I\mathbf{h}_I^H\right)^2} + \eta\sigma^2$ | $\frac{\sqrt{\gamma_0\sigma^2}}{\left(\mathbf{h}_I\mathbf{h}_I^H\right)} \cdot \mathbf{h}_I^H$ | $\mathbf{0}$ | $\frac{\gamma_0\sigma^2}{\left(\mathbf{h}_I\mathbf{h}_I^H\right)}$ |
| $\frac{\gamma_0\eta\sigma^2\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right)}{\left(\mathbf{h}_I\mathbf{h}_I^H\right)^2} + \eta\sigma^2 < E_0 < \frac{\gamma_0\eta\sigma^2\left(\mathbf{h}_E\mathbf{h}_E^H\right)^2}{\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right)} + \eta\sigma^2$ | $\alpha \cdot \mathbf{h}_I^H + \beta \cdot \mathbf{h}_\perp^H$ | $\mathbf{0}$ | $\lambda_I\sigma^2 + \lambda_E\left(\frac{E_0}{\eta} - \sigma^2\right)$ |
| $E_0 \geq \frac{\gamma_0\eta\sigma^2\left(\mathbf{h}_E\mathbf{h}_E^H\right)^2}{\left(\mathbf{h}_I\mathbf{h}_E^H\mathbf{h}_E\mathbf{h}_I^H\right)} + \eta\sigma^2$ | $\frac{\sqrt{\frac{E_0-\eta\sigma^2}{\eta(1+c^2)}}}{\left(\mathbf{h}_E\mathbf{h}_E^H\right)} \cdot \mathbf{h}_E^H$ | $c \cdot \mathbf{w}_I^*$ | $\frac{E_0-\eta\sigma^2}{\eta\left(\mathbf{h}_E\mathbf{h}_E^H\right)}$ |

**Table 7.1:** Optimal precoders and the corresponding transmit power required for the case of one ID and one EH receiver

### 7.2.3.2 The case of $K$ EH Receivers

We proceed to extend our analyses to multiple EH receivers, and show that the optimal precoding scheme will be to employ the information precoder only, as detailed below.

**Only SINR Constraint is Active:** This case is similar to the case where there is only one EH receiver, and the optimal precoding scheme will be to employ the information precoder only, which is obtained in (7.28). The harvested energy requirement should satisfy

$$E_0 \leq \min_k \left[\eta\mathbf{h}_E^k\mathbf{w}_I\mathbf{w}_I^H\left(\mathbf{h}_E^k\right)^H + \eta\sigma^2\right] \Rightarrow E_0 \leq \min_k \left\{\frac{\gamma_0\eta\sigma^2\left[\mathbf{h}_I\left(\mathbf{h}_E^k\right)^H\mathbf{h}_E^k\mathbf{h}_I^H\right]}{\left(\mathbf{h}_I\mathbf{h}_I^H\right)^2} + \eta\sigma^2\right\}. \tag{7.64}$$

**Only Harvested Energy Constraints are Active:** This corresponds to the case where the harvested energy requirement is high, and we can obtain $\lambda_I = 0$, since the SINR constraint is not active. It is easy to observe from the stationarity conditions that each $\mathbf{w}_E^i$ is parallel to $\mathbf{w}_I$, and therefore we can express each $\mathbf{w}_E^i$ as

$$\mathbf{w}_E^i = c_i \cdot \mathbf{w}_I, \ \forall i \in \mathscr{K}. \tag{7.65}$$

With the SINR constraint being over-satisfied, we can further obtain $\sum_{i=1}^{K} c_i^2 < \frac{1}{\gamma_0}$ by following a similar approach in (7.46). Accordingly, the power ratio $\rho$ defined in (7.62) for the case of $K > 1$ EH receivers can be further obtained

as

$$\rho = \frac{\mathbf{w}_I^H \mathbf{w}_I}{\mathbf{w}_I^H \mathbf{w}_I + \sum_{i=1}^{K} c_i^2 \mathbf{w}_I^H \mathbf{w}_I} = \frac{1}{1 + \sum_{i=1}^{K} c_i^2} \;\Rightarrow\; \rho > \frac{1}{1+\frac{1}{\gamma_0}} = \frac{\gamma_0}{1+\gamma_0}. \tag{7.66}$$

Note that by setting $c_i = 0, \forall i \in \mathscr{K}$, we obtain $\rho = 1$, which means that the optimality is achieved by employing the information precoder only.

**Both Constraints are Active:** In this case, similar to the derivation in (7.55)-(7.57), we can obtain the total transmit power as

$$P_{TX} = \lambda_I \sigma^2 + \sum_{k=1}^{K} \lambda_E^k \left( \frac{E_0}{\eta} - \sigma^2 \right). \tag{7.67}$$

Following a similar approach in **Proposition 3**, it is obtained by contradiction that the optimality is achieved by employing only the information precoder, and $\left(\mathbf{w}_E^i\right)^* = \mathbf{0}, \forall i \in \mathscr{K}$.

With the above mathematical analyses, we can conclude that the optimality for the considered scenario is to employ the information precoder $\mathbf{w}_I$ only. This can also be observed based on the fact that EH receivers do not need to decode the symbols, and therefore energy precoding is indeed not necessary. Our contribution here is that we mathematically prove the above observation and obtain the closed-form expressions for the special case of $K = 1$ EH receiver with the above analyses.

### 7.2.3.3 SDR Approach to Solve $\mathscr{P}_{7.1}$

Based on the above analyses and by introducing $\mathbf{W}_I = \mathbf{w}_I \mathbf{w}_I^H$, $\mathbf{D}_I = \mathbf{h}_I^H \mathbf{h}_I$ and $\mathbf{D}_E^k = \left(\mathbf{h}_E^k\right)^H \mathbf{h}_E^k$, $\mathscr{P}_{7.1}$ can be simplified and further transformed into a SDP, given by

$$\begin{aligned} \mathscr{P}_{7.3}: \; & \min_{\mathbf{W}_I} \; tr\{\mathbf{W}_I\} \\ s.t. \quad & \frac{1}{\gamma_0} tr\{\mathbf{D}_I \mathbf{W}_I\} \geq \sigma^2 \\ & tr\left\{\mathbf{D}_E^k \mathbf{W}_I\right\} + \sigma^2 \geq \frac{E_0}{\eta}, \; \forall k \in \mathscr{K} \\ & \mathbf{W}_I \succeq 0, \; \text{rank}\{\mathbf{W}_I\} = 1 \end{aligned} \tag{7.68}$$

By dropping the rank-1 constraint for $\mathbf{W}_I$ in $\mathscr{P}_{7.3}$, the relaxed optimisation problem becomes convex, and can be efficiently solved by convex optimisation tools. When $\text{rank}\{\mathbf{W}_I^*\} = 1$, the optimal solution to $\mathscr{P}_{7.1}$ can be obtained by employing the eigenvalue decomposition of $\mathbf{W}_I^*$, given by

$$\mathbf{w}_I^* = \mathbf{U}\boldsymbol{\Sigma}^{1/2}, \tag{7.69}$$

where $\mathbf{U}$ and $\boldsymbol{\Sigma}$ correspond to the eigenvectors and eigenvalues of $\mathbf{W}_I^*$ respectively, and we have $\mathbf{W}_I^* = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^H$. It has been shown in [126, 127] that the solution for the relaxed version of $\mathscr{P}_{7.3}$ satisfies

$$\text{rank}\,(\mathbf{W}_I^*) \leq \sqrt{K+1}. \tag{7.70}$$

When there are no more than 2 EH receivers in the system ($K \leq 2$), explicitly we have $\sqrt{K+1} < 2$ and the obtained solution is guaranteed to be rank-1, which means that in this case the semi-definite relaxation (SDR) is not a relaxation but an optimal solution to the original problem. When $K > 2$, while the obtained solution cannot be guaranteed to be rank-1, we show below in Table 7.2 that in most cases the obtained solutions still satisfy the rank-1 constraint and are therefore optimal, when the number of EH receivers $K$ is small. The results in Table 7.2 are obtained based on 5000 channel realisations with $N_t = 12$, $\eta = 0.35$, $E_0 = 5\text{mW}$ and $\gamma_0 = 10\text{dB}$. On the other hand, when the obtained rank of $\mathbf{W}_I^*$ is larger than 1, we can obtain a feasible close-to-optimal solution as

$$\mathbf{w}_I = \tau \cdot \mathbf{w}_I^* = \tau \cdot \mathbf{U}\boldsymbol{\Sigma}^{1/2}, \tag{7.71}$$

where $\tau \geq 1$ and can be obtained as

$$\tau = \max\left\{\sqrt{\frac{\gamma_I\sigma^2}{\left|\mathbf{h}_I\mathbf{U}\boldsymbol{\Sigma}^{1/2}\right|^2}}, \sqrt{\frac{E_1 - \eta\sigma^2}{\eta\left|\mathbf{h}_E^1\mathbf{U}\boldsymbol{\Sigma}^{1/2}\right|^2}}, \ldots, \sqrt{\frac{E_K - \eta\sigma^2}{\eta\left|\mathbf{h}_E^K\mathbf{U}\boldsymbol{\Sigma}^{1/2}\right|^2}}\right\}, \tag{7.72}$$

| Number of Users $(1+K)$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Average rank of $\mathbf{W}_I^*$ | 1 | 1 | 1 | 1.0528 | 1.1736 |
| Maximum rank of $\mathbf{W}_I^*$ | 1 | 1 | 1 | 2 | 2 |
| Rank-1 percentage of $\mathbf{W}_I^*$ | 100% | 100% | 100% | 94.72% | 82.64% |

**Table 7.2:** Average rank, maximum rank and rank-1 percentage of $\mathbf{W}_I^*$ for $\mathscr{P}_{7.3}$, $N_t = 12$, $\eta = 0.35$, $E_0 = 5$mW, $\gamma_0 = 10$dB

which guarantees that all the constraints in $\mathscr{P}_{7.1}$ are met. For the small-scale MIMO system considered in the SWIPT technique in Chapter 7.2, the SDR approach can be effectively applied to obtain the solution of $\mathscr{P}_{7.1}$ and the optimisation problems in the subsequent chapters.

### 7.2.4 Low-Complexity Hybrid Precoding via SVD

The FD MIMO SWIPT precoding discussed in Chapter 7.2.3 requires a dedicated RF chain for each antenna element, which is inefficient in both the hardware complexity and power consumption. Towards energy-efficient SWIPT, we proceed to study the hybrid structure where the BS employs $N_{RF}^t$ $(1 \leq N_{RF}^t < N_t)$ RF chains, and the precoding is divided into the analog domain and the digital domain. Accordingly, we can express the transmit signal vector as

$$\mathbf{x} = \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s}, \tag{7.73}$$

where $\mathbf{s} = \left[s_I, s_E^1, s_E^2, ..., s_E^K\right]^T$ denotes the data symbol vector, and $\mathbf{s} \in \mathbb{C}^{(K+1)\times 1}$. $\mathbf{F}_{RF} \in \mathbb{C}^{N_t \times N_{RF}^t}$ denotes the analog precoder implemented with phase shifter networks. $\mathbf{F}_{BB} \in \mathbb{C}^{N_{RF}^t \times (K+1)}$ represents the digital precoder and we decompose it into

$$\mathbf{F}_{BB} = \left[\mathbf{f}_I, \mathbf{f}_E^1, \mathbf{f}_E^2, ..., \mathbf{f}_E^K\right]. \tag{7.74}$$

We consider both a fully-connected structure and a partially-connected structure as in Chapter 7.1.2 and Fig. 2.9, where $M = N_t/N_{RF}^t$ denotes the number of antennas connected to each RF chain.

In this section, a low-complexity hybrid precoding method for MIMO SWIPT is proposed based on SVD, which serves as a benchmark to be com-

pared with the proposed iterative method in Chapter 7.2.5. When the hybrid precoding is employed, it is generally difficult to directly solve the joint optimisation problem, due to the non-convex constant modulus constraint of the analog precoder that is in the form of (2.64)-(2.67). Therefore, to remove the non-convex constant modulus constraint in the optimisation problem, it is intuitive that we first design the analog precoder $\mathbf{F}_{RF}$, followed by the design of the low-dimensional digital precoder $\mathbf{F}_{BB}$ with convex optimisation. To be specific, we express the SVD of the channel as

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{7.75}$$

where $\mathbf{U}$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_{N_t}]$ are the left- and right-singular vectors. Each phase $\varphi_{m,n}$ of the analog precoder $\mathbf{F}_{RF}$ is selected as

$$\varphi_{m,n} = \theta_{m,n}, \ \ m \in \mathscr{N}, \ \ n \in \mathscr{U}, \tag{7.76}$$

where $\mathscr{U} = \{1, 2, ..., N_{RF}^t\}$ and $\theta_{m,n}$ is the phase of the $m$-th element in $\mathbf{v}_n$. While we employ an analog precoding design based on SVD, other channel-dependent analog designs can also be applied. With $\mathbf{F}_{RF}$ obtained, the optimisation problem to obtain $\mathbf{F}_{BB}$ can be formulated as

$$\begin{aligned} \mathscr{P}_{7.4}: \ & \min_{\mathbf{f}_I, \mathbf{f}_E^i} \ \|\mathbf{F}_{RF}\mathbf{f}_I\|^2 + \sum_{i=1}^{K} \left\|\mathbf{F}_{RF}\mathbf{f}_E^i\right\|^2 \\ s.t. \quad & \gamma_I \geq \gamma_0 \\ & E_k \geq E_0, \ \ \forall k \in \mathscr{K} \end{aligned} \tag{7.77}$$

While the analyses in Chapter 7.2.3 is conducted for the FD precoding approaches, similar analyses can be performed for $\mathscr{P}_{7.4}$ of the hybrid precoding, since in such case the analog precoder $\mathbf{F}_{RF}$ can be regarded as a fixed matrix and further incorporated into the channel. A similar conclusion can be drawn that the optimality for $\mathscr{P}_{7.4}$ is to employ the low-dimensional information precoder only, and therefore the SDP form of $\mathscr{P}_{7.4}$ can be simplified

into

$$\begin{aligned}\mathscr{P}_{7.5}: \ & \min_{\mathbf{F}_I} \ tr\left\{\mathbf{F}_{RF}\mathbf{F}_I\mathbf{F}_{RF}^H\right\} \\ s.t. \quad & tr\left\{\mathbf{D}_I\mathbf{F}_{RF}\mathbf{F}_I\mathbf{F}_{RF}^H\right\} \geq \gamma_0\sigma^2 \\ & tr\left\{\mathbf{D}_E^k\mathbf{F}_{RF}\mathbf{F}_I\mathbf{F}_{RF}^H\right\} \geq \frac{E_0}{\eta} - \sigma^2, \ \forall k \in \mathscr{K} \\ & \mathbf{F}_I \succeq 0, \ \text{rank}\{\mathbf{F}_I\} = 1\end{aligned} \tag{7.78}$$

where $\mathbf{F}_I = \mathbf{f}_I\mathbf{f}_I^H$. By dropping the rank-1 constraint, $\mathscr{P}_{7.5}$ can also be effectively solved. As the rank of the obtained $\mathbf{F}_I$ is only related to the number of users $(1+K)$, the rank result and approach in Chapter 7.2.3.3 can be trivially extended to the hybrid case of $\mathscr{P}_{7.5}$.

**Extension to Partially-Connected Structures:** When partially-connected structures are considered, each $\mathbf{f}_k^p \in \mathbb{C}^{M\times 1}$. Similar to the designs for partially-connected structures in Chapter 7.1, each entry of $\mathbf{f}_k^p$ can be obtained as

$$\left[\mathbf{f}_k^p\right]_m = \frac{1}{M}e^{j\theta_{n,k}}, \ k = (n-1)M + m. \tag{7.79}$$

where $\theta_{n,k}$ denotes the phase of the $k$-th entry in $\mathbf{v}_n$. While the partially-connected structures require an increased transmit power $P_{TX}$ to achieve the same performance requirements compared to the fully-connected structures, we further note that the total power consumption $P_{BS}$ for partially-connected structures will in fact be much lower than the FD case and fully-connected structures, due to the reduced number of RF chains and phase shifters required. We shall quantify this favourable tradeoff in terms of the total power consumption at the BS in Chapter 7.2.7.

### 7.2.5 An Iterative Approach via a Geometrical Representation

Based on the results given in Table 7.1 and the analyses in Chapter 7.2.3, it is observed that the optimal precoding strategy is to employ the information precoder $\mathbf{w}_I$ only (by noting that each $c_i$ can be equal to 0). By considering the effective channel expression, similar results can be derived for the hybrid

case. These two observations lead to

$$\mathbf{W} = \left[\mathbf{w}_I, \mathbf{0}^{N_t \times K}\right], \ \mathbf{F}_{BB} = \left[\mathbf{f}_I, \mathbf{0}^{(K+1)\times K}\right]. \tag{7.80}$$

Accordingly, we propose an iterative approach where we alternately update the analog precoder and the digital precoder. To be more specific, for the design of the analog precoder $\mathbf{F}_{RF}$, instead of employing the SVD, we propose to minimise the Euclidean distance between the optimal FD precoder and the hybrid precoder, which can be formulated as

$$\begin{aligned} \mathscr{P}_{7.6}: \ & \min_{\mathbf{F}_{RF}} \ \|\mathbf{W} - \mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 \\ & s.t. \quad \mathbf{F}_{RF} \in \mathscr{F} \end{aligned} \tag{7.81}$$

where we denote $\mathscr{F}$ as the set that consists of the matrices that satisfy the constant modulus constraint for each of their entries. Based on (7.80), the objective function of $\mathscr{P}_{7.6}$ can be further decomposed into

$$\begin{aligned} & \|\mathbf{W} - \mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 \\ = & \left\| \left[\mathbf{w}_I, \mathbf{0}^{N_t \times K}\right] - \mathbf{F}_{RF} \left[\mathbf{f}_I, \mathbf{0}^{(K+1)\times K}\right] \right\|_F^2 \\ = & \|\mathbf{w}_I - \mathbf{F}_{RF}\mathbf{f}_I\|_F^2 \\ = & \sum_{i=1}^{N_t} \left| [\mathbf{w}_I]_i - \mathbf{f}_{RF}^i \mathbf{f}_I \right|^2, \end{aligned} \tag{7.82}$$

where we note that each $[\mathbf{w}_I]_i$ is a scalar. In (7.82), we decompose $\mathbf{F}_{RF} = \left[\left(\mathbf{f}_{RF}^1\right)^T, \left(\mathbf{f}_{RF}^2\right)^T, ..., \left(\mathbf{f}_{RF}^{N_t}\right)^T\right]^T$ and $\mathbf{f}_{RF}^i \in \mathbb{C}^{1\times(K+1)}$ is the $i$-th row of $\mathbf{F}_{RF}$. As can be observed, the objective function is decomposed into $N_t$ independent sub-functions by row, and therefore the optimisation problem $\mathscr{P}_{7.6}$ is equivalent to minimising each of the $N_t$ independent sub-problems. Accordingly, we

formulate the $i$-th sub-problem as

$$\begin{aligned} \mathscr{P}_{7.7}^{i}: \ & \min_{\mathbf{f}_{RF}^{i}} \left| [\mathbf{w}_I]_i - \mathbf{f}_{RF}^{i} \mathbf{f}_I \right| \\ s.t. \quad & \mathbf{f}_{RF}^{i} \in \mathscr{G} \end{aligned} \tag{7.83}$$

where $\mathscr{G}$ denotes the set of row vectors that satisfy the constant modulus constraint for each of their entries. For simplicity we introduce

$$[\mathbf{t}]_m = \left[\mathbf{f}_{RF}^{i}\right]_m [\mathbf{f}_I]_m, \tag{7.84}$$

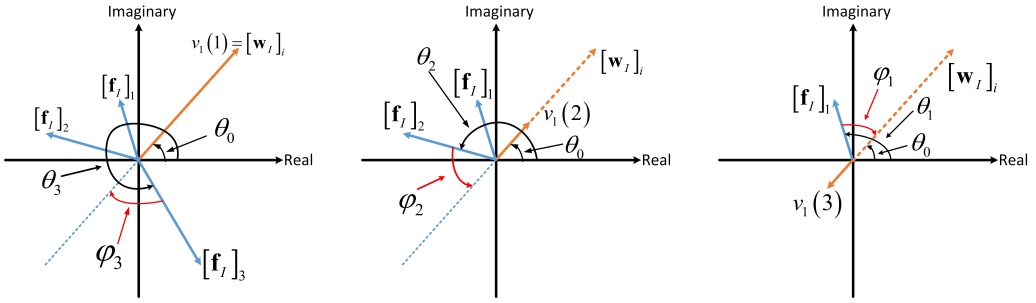
and the optimisation problem can be further transformed into

$$\begin{aligned} \mathscr{P}_{7.8}^{i}: \ & \min_{\mathbf{f}_{RF}^{i}} |u_i| \\ s.t. \quad & u_i = [\mathbf{w}_I]_i - \sum_{m=1}^{K+1} [\mathbf{t}]_m \\ & \mathbf{f}_{RF}^{i} \in \mathscr{G} \end{aligned} \tag{7.85}$$

#### 7.2.5.1 Optimal Analog Precoder

As each entry of $\mathbf{f}_{RF}^{i}$ is of constant modulus, therefore the multiplication of each $\left[\mathbf{f}_{RF}^{i}\right]_m$ to the corresponding $[\mathbf{f}_I]_m$ in (7.84) is equivalent to an angle rotation in the complex plane. Moreover, since $[\mathbf{w}_I]_i$ in $\mathscr{P}_{7.8}^{i}$ is a scalar, we can therefore employ a geometric representation to arrive at an optimal solution efficiently. An explanatory geometrical representation for the case of one ID and $K = 2$ EH receivers is shown in Fig. 7.6, where the dashed brown arrow represents the optimal FD solution $[\mathbf{w}_I]_i$, the solid brown arrow denotes the value of the function $v_i(m)$ that is to be introduced in (7.87), and the solid blue arrows denote each entry in $\mathbf{f}_I$. We denote $\theta_0$ as the phase of $[\mathbf{w}_I]_i$, $\theta_m$ as the phase of $[\mathbf{f}_I]_m$, and we assume $\theta_0$, $\theta_m \in [0, 2\pi)$, as shown in Fig. 7.6.

Geometrically, it is observed that the optimal solution for $\mathbf{f}_{RF}^{i}$ that minimises $|u_i|$ in $\mathscr{P}_{7.8}^{i}$ is to rotate each $[\mathbf{f}_I]_m$ such that each resulting $[\mathbf{t}]_m$ is collinear to $[\mathbf{w}_I]_i$. Accordingly, we introduce the algorithm employed to solve the sub-




**(a)** Step 1: Rotate $[\mathbf{f}_I]_3$ to the opposite side of $[\mathbf{w}_I]_i$ **(b)** Step 2: Rotate $[\mathbf{f}_I]_2$ to the opposite side of $[\mathbf{w}_I]_i$ **(c)** Step 3: Rotate $[\mathbf{f}_I]_1$ to the same side of $[\mathbf{w}_I]_i$

**Figure 7.6:** Geometric interpretation of $\mathscr{P}_{7.9}^i$ and the solution for the example of one ID and $K = 2$ EH receivers

problem $\mathscr{P}_{7.8}^i$ based on successive phase rotation. To be specific, to achieve collinearity for each resulting $[\mathbf{t}]_m$, the phase of the corresponding $\left[\mathbf{f}_{RF}^i\right]_m$ can be obtained based on Fig. 7.6 as

$$\varphi_m = \theta_0 - \theta_m, \text{ or } \varphi_m = \theta_0 + \pi - \theta_m, \tag{7.86}$$

which is dependent on whether $[\mathbf{f}_I]_m$ is rotated to the same direction of $[\mathbf{w}_I]_i$ or the opposite direction of $[\mathbf{w}_I]_i$. Define a function $v_i$ with respect to $m$ that represents the difference between the optimal FD precoder and the sum of the previously rotated components of $\mathbf{f}_I$ as

$$\textbf{Function}: v_i(m) = [\mathbf{w}_I]_i - \sum_{j=1}^{m-1} [\mathbf{t}]_j, \tag{7.87}$$

and we further define $v_i(1) = [\mathbf{w}_I]_i$. $v_i(m)$ therefore represents the residual portion of $[\mathbf{w}_I]_i$ to be cancelled, and it is observed that the value of each $\varphi_m$ is dependent on $v_i(m)$. More specifically, from the definition of $v_i(m)$, each $\left[\mathbf{f}_{RF}^i\right]_m$ should be rotated to the opposite side of $v_i(m)$. Moreover, to guarantee optimum and that the resulting objective function $|u_i|$ in $\mathscr{P}_{7.8}^i$ is minimised, we should sort the elements in $\mathbf{f}_I$ in the descending order of amplitude, and the elements with larger amplitudes should be rotated first. An illustrative example is given in Fig. 7.6, where $[\mathbf{f}_I]_3$ is firstly rotated, followed

by the rotation of $[\mathbf{f}_I]_2$. $[\mathbf{f}_I]_1$ has the smallest amplitude and is therefore rotated at last. We summarise the above algorithm to obtain $\mathbf{f}_{RF}^i$ in Algorithm 7.4.

**Algorithm 7.4** Optimal Analog Precoding Solution for $\mathscr{P}_{7.8}^i$

**input** : $[\mathbf{w}_I]_i$, $\mathbf{f}_I$
**output** : $\mathbf{f}_{RF}^i$
$\theta_0 = \arg\{[\mathbf{w}_I]_i\}$, $\hat{\mathbf{f}}_I = \text{sort}(\mathbf{f}_I, \text{d})$;
**for** $m = 1 : N_{RF}^t$ **do**
  Obtain $m_0 = \text{find}\left\{\left[\hat{\mathbf{f}}_I\right]_m = \mathbf{f}_I\right\}$, $\theta_{m_0} = \arg\left\{[\mathbf{f}_I]_{m_0}\right\}$;
  Calculate $[v_i]_m$;
  **if** $\arg\{[v_i]_m\} = \arg\{[\mathbf{w}_I]_i\}$ **then**
    $\varphi_{m_0} = \theta_0 + \pi - \theta_{m_0}$;
  **end if**
  **if** $\arg\{[v_i]_m\} = \pi + \arg\{[\mathbf{w}_I]_i\}$ **then**
    $\varphi_{m_0} = \theta_0 - \theta_{m_0}$;
  **end if**
  $\left[\mathbf{f}_{RF}^i\right]_{m_0} = \frac{1}{N_t} e^{j\varphi_{m_0}}$.
**end for**

In Algorithm 7.4, $\hat{\mathbf{a}} = \text{sort}(\mathbf{a}, \text{d})$ denotes the function that sorts the elements of $\mathbf{a}$ in a descending order of amplitude, and the re-ordered vector is denoted as $\hat{\mathbf{a}}$. The function $x = \text{find}\{a = \mathbf{b}\}$ means that $[\mathbf{b}]_x = a$, and we denote $\arg\{a\}$ as the phase of $a$. With the above algorithm, the optimal $\mathbf{f}_{RF}^i$ can be efficiently obtained and the resulting $|u_i|$ is guaranteed to be the minimal. We perform Algorithm 7.4 to calculate each $\mathbf{f}_{RF}^i$ for $N_t$ times, and the optimal analog precoder $\mathbf{F}_{RF}$ can be obtained.

### 7.2.5.2 Iterative Algorithm

It is observed that the design of each $\mathbf{f}_{RF}^i$ in Algorithm 7.4 requires the knowledge of the digital precoder $\mathbf{f}_I$, and accordingly we propose an iterative design where we alternately update $\mathbf{F}_{RF}$ and $\mathbf{F}_{BB}$ until convergence or a maximum number of iterations is reached. The proposed algorithm is summarised in Algorithm 7.5, where $\mathbf{w}_I$ is the optimal FD precoder for $\mathscr{P}_{7.3}$ in Chapter 7.2.3, $\mathbf{F}_{BB}^0$ is the initial low-dimensional digital precoder of the hybrid precoding obtained from $\mathscr{P}_{7.5}$ in Chapter 7.2.4, and $N_{max}$ denotes the

maximum iteration number. We introduce $\Delta$ as a variable that represents the convergence accuracy, and $\Delta_{th}$ denotes the accuracy threshold.

**Algorithm 7.5** The Iterative Hybrid Precoding Design

**input** : $\mathbf{w}_I$, $\mathbf{F}_{BB}^0$
**output** : $\mathbf{F}_{RF}^*$, $\mathbf{F}_{BB}^*$
$n = 0$, $\mathbf{W}_h^{(0)} = \mathbf{0}$;
**while** $n \leq N_{max}$ and $\Delta \geq \Delta_{th}$ **do**
  Obtain $\mathbf{f}_I^{(n)}$ from $\mathbf{F}_{BB}^{(n)}$ based on (7.74);
  Obtain $\mathbf{F}_{RF}^{(n+1)}$ by Algorithm 7.4 with $\mathbf{f}_I^{(n)}$;
  Obtain $\mathbf{F}_{BB}^{(n+1)}$ by solving $\mathscr{P}_{7.5}$ with $\mathbf{F}_{RF}^{(n+1)}$;
  $\mathbf{W}_h^{(n+1)} = \mathbf{F}_{RF}^{(n+1)}\mathbf{F}_{BB}^{(n+1)}$;
  $\Delta = \left\| \mathbf{W}_h^{(n+1)} - \mathbf{W}_h^{(n)} \right\|$
  $n = n + 1$;
**end while**
$\mathbf{F}_{RF}^* = \mathbf{F}_{RF}^{(n)}$, $\mathbf{F}_{BB}^* = \mathbf{F}_{BB}^{(n)}$.

**Convergence Discussion:** It has been shown in Table 7.2 that for the considered small-scale multiple-antenna system, in most cases the SDR approach can obtain the optimal rank-1 solution. In this case, since the subproblems to obtain $\mathbf{F}_{RF}$ and $\mathbf{F}_{BB}$ in each iteration are solved optimally, the iterative design in Algorithm 7.5 is guaranteed to converge [69, 121]. Nevertheless, when the rank of the obtained solution is larger than 1, while the convergence cannot be explicitly proven, a feasible close-to-optimal solution can be obtained based on (7.71), (7.72), and it is observed in our simulations that the proposed iterative algorithm is also shown to be convergent. Furthermore, our approach also includes a maximum number of iterations $N_{max}$ to terminate the iterations and return a solution.

**Extension to Partially-Connected Structures:** When partially-connected structures are considered for the proposed iterative design, each $\mathbf{f}_{RF}^i$ in (7.82) only has one non-zero element, and we only need to rotate this entry to the opposite side of $[\mathbf{w}_I]_i$, which greatly simplifies the design. We summarise the corresponding algorithm to obtain $\mathbf{F}_{RF}$ for partially-connected structures in Algorithm 7.6, where the function $\lceil x \rceil$ denotes the minimum integer that is

not smaller than $x$.

**Algorithm 7.6** Analog Precoder of the Iterative Scheme for Partially-Connected Structures

**input** : $\mathbf{w}_I$, $\mathbf{f}_I$
**output** : $\mathbf{F}_{RF}$
**for** $i = 1 : N_t$ **do**
  $\mathbf{f}_{RF}^i = \mathbf{0}^{1\times(K+1)}$;
  Obtain $[\mathbf{w}_I]_i$, then $\theta_0 = \arg\{[\mathbf{w}_I]_i\}$;
  Calculate $m_0 = \left\lceil \frac{i}{M} \right\rceil$;
  $\theta_{m_0} = \arg\left\{[\mathbf{f}_I]_{m_0}\right\}$, $\phi_{m_0} = \theta_0 + \pi - \theta_{m_0}$;
  $\left[\mathbf{f}_{RF}^i\right]_{m_0} = e^{j\phi_{m_0}}$;
**end for**
$\mathbf{F}_{RF} = \left[\left(\mathbf{f}_{RF}^1\right)^T, \left(\mathbf{f}_{RF}^2\right)^T, ..., \left(\mathbf{f}_{RF}^{N_t}\right)^T\right]^T$.

### 7.2.6 Computational Cost Analysis

For both proposed techniques, it is observed that the dominant complexity arises from solving the optimisation problems. For each optimisation problem, based on [167, 168] the interior-point algorithm for solving the dual problem of an $M$-dimensional optimisation with $N$ variables requires $\mathscr{O}\left\{\sqrt{NM}\left(N^3M^2 + N^2M^3\right)\right\}$ floating-point operations. For the FD method based on the optimisation problem $\mathscr{P}_{7.3}$, $\mathbf{W}_I$ and each $\mathbf{W}_E^i$ are $N_t$-dimensional, and there is one ID and $K$ EH receivers in the system, which leads to $M = N_t$, and $N = K + 1$. For the proposed hybrid precoding based on SVD, we can similarly obtain that $M = N = K + 1$, and the computational cost of the proposed iterative scheme will be $N_{max}$ times higher than that of the hybrid precoding based on SVD, since the SDP needs to be performed $N_{max}$ times, as observed in Algorithm 7.5. Accordingly, we can obtain the computational cost of each scheme in terms of the number of floating-point operations, which can be expressed as

$$C_{Fully} = \mathscr{O}\left\{\sqrt{(K+1)N_t}\left[(K+1)^3N_t^2 + (K+1)^2N_t^3\right]\right\}, \tag{7.88}$$

$$C_{Hybrid}^{SVD} = \mathscr{O}\left\{\sqrt{(K+1)(K+1)}\left[2(K+1)^3(K+1)^2\right]\right\} = \mathscr{O}\left\{2(K+1)^6\right\}, \tag{7.89}$$

| Schemes | Number of UEs | | |
|---|---|---|---|
| | 1 ID, 3 EHs | 1 ID, 4 EHs | 1 ID, 5 EHs |
| Fully-digital | $\mathscr{O}\left(2.6\times10^5\right)$ | $\mathscr{O}\left(4.7\times10^5\right)$ | $\mathscr{O}\left(7.9\times10^5\right)$ |
| Hybrid, SVD | $\mathscr{O}\left(8.1\times10^3\right)$ | $\mathscr{O}\left(3.1\times10^4\right)$ | $\mathscr{O}\left(9.3\times10^4\right)$ |
| Hybrid, Iterative | $\mathscr{O}\left(3.2\times10^4\right)$ | $\mathscr{O}\left(1.2\times10^5\right)$ | $\mathscr{O}\left(3.7\times10^5\right)$ |

**Table 7.3:** Computational complexity of the fully-digital scheme and the hybrid schemes for $N_t = 12$, $N_{max} = 4$

$$C_{Hybrid}^{Iterative} = \mathscr{O}\left\{N_{max}\left(K+1\right)\left[2(K+1)^3(K+1)^2\right]\right\} = \mathscr{O}\left\{2N_{max}(K+1)^6\right\}. \tag{7.90}$$

It is observed that the computational costs of both hybrid techniques are irrelevant to the number of transmit antennas $N_t$ at the BS, and the complexity-reduction gain will be higher when $N_t$ increases. The computational cost of each method in a representative MIMO SWIPT scenario is explored in Table 7.3 below, where the number of transmit antennas at the BS is $N_t = 12$. It is observed that compared to the FD case, both of the proposed hybrid precoding approaches require less computational costs, and the hybrid approach based on SVD is the most computationally efficient one.

### 7.2.7 Power Consumption Model

To demonstrate the significant power savings introduced by the hybrid analog-digital architectures, we introduce the power consumption model employed in the simulations before presenting the numerical results. For the FD case, the analog phase shifters are not needed. Accordingly, based on [37, 81] the power consumption model at the BS is given by

$$\begin{aligned}
P_{BS}^{FD} &= N_t\left(N_t+1\right)P_{PA} + P_{BB} + N_t\left(P_{RFC}+P_{DAC}\right),\\
P_{BS}^{Fully} &= N_t\left(N_{RF}^t+1\right)P_{PA} + N_tN_{RF}^tP_{PS} + P_{BB} + N_{RF}^t\left(P_{RFC}+P_{DAC}\right),\\
P_{BS}^{Partially} &= N_tP_{PA} + N_tP_{PS} + P_{BB} + N_{RF}^t\left(P_{RFC}+P_{DAC}\right),
\end{aligned} \tag{7.91}$$

where $P_{BS}^{FD}$, $P_{BS}^{Fully}$, and $P_{BS}^{Partially}$ denote the total power consumption at the BS for the FD case, hybrid fully-connected and hybrid partially-connected structures, respectively. In (7.91), $P_{PA} = (1/\eta_0)P_{TX}$ is the power consumed at

the power amplifier to generate the transmit power $P_{TX}$, with $\eta_0$ being the power amplifier efficiency. $P_{PS}$ represents the power consumption for phase shifters, $P_{RFC}$ the power consumption for the RF chains, $P_{DAC}$ the power consumption for the DACs, and $P_{BB}$ the power consumption for the baseband processing.

### 7.2.8 Numerical Results

In this section, we conduct Monte Carlo simulations to evaluate the performance of the proposed hybrid schemes. There are a total number of $N_t = 12$ transmit antennas at the BS, and we consider one ID and $K = 3$ EH receivers in the system. The typical values of the power consumption for each hardware component follow [81]. The simulation parameters are summarised in Table 7.4, and remain constant throughout the simulations unless otherwise stated.

| Simulation Parameters | Values | Simulation Parameters | Values |
|---|---|---|---|
| Antenna Number at the BS, $N_t$ | 12 | Distance of EH receivers and BS, $D_E^k$ (m) | 5 |
| Number of ID | 1 | Channel noise power, $\sigma^2$ (mW) | 0.1 |
| Number of EH receivers, $K$ | 3 | Energy transfer efficiency, $\eta$ | 0.35 |
| Number of RF chains, $N_{RF}^t$ | 4 | Power amplifier efficiency, $\eta_0$ | 0.5 |
| Propagation constant, $\alpha_0$ | 1 | Power of phase shifters, $P_{PS}$ (mW) | 30 |
| Pathloss coefficient, $\beta$ | 2 | Power of RF chains, $P_{RFC}$ (mW) | 40 |
| Shadow fading, $C_I$, $C_E^k$ | 1 | Power of DAC, $P_{DAC}$ (mW) | 200 |
| Distance of ID and BS, $D_I$ (m) | 10 | Power of baseband processing, $P_{BB}$ (mW) | 5 |

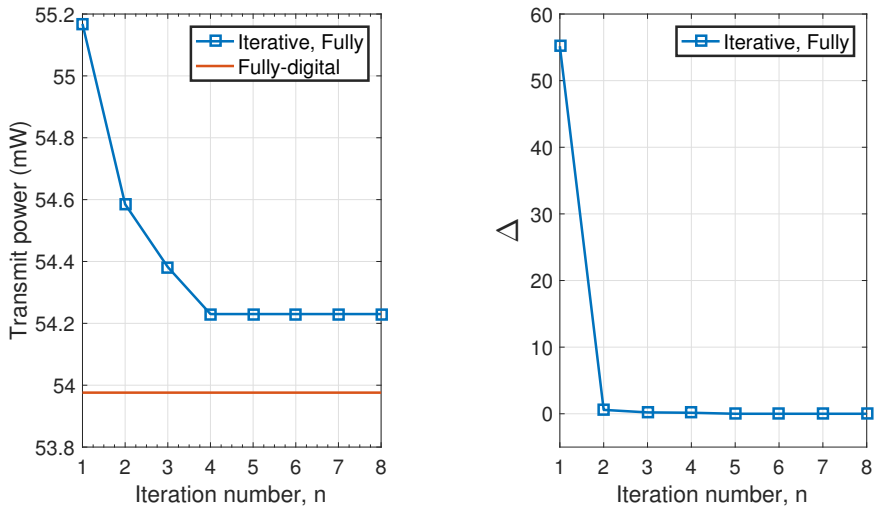
**Table 7.4:** Simulation Parameters

We compare our proposed schemes with the FD method, and the following abbreviations are applied for clarity:

- "Fully-digital": conventional FD approach at the BS, $\mathscr{P}_{7.3}$;
- "SVD, Fully/Partially": the proposed hybrid precoding based on the SVD in Chapter 7.2.4;
- "Iterative, Fully/Partially": the proposed iterative hybrid design by Algorithm 7.5 in Chapter 7.2.5.

In Fig. 7.7, we evaluate the convergence of the proposed iterative algorithm by plotting the value of the transmit power and the value of $\Delta$ with

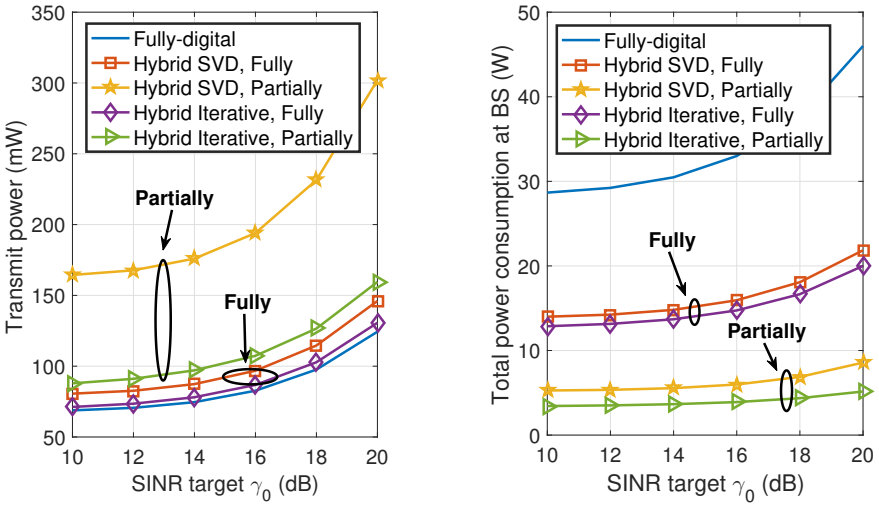
respect to the iteration number $n$. We select $\mathbf{F}_{BB}^{0}$ as the digital precoder obtained by $\mathscr{P}_{7.5}$ in Chapter 7.2.4. It can be observed that the proposed iterative scheme is convergent within $n = 4$ iterations. Furthermore, the performance gap compared to the FD case is marginal.




**(a)** Transmit power $P_{TX}$ (mW) **(b)** Value of $\Delta$

**Figure 7.7:** Convergence of the iterative algorithm for $N_t = 12$, one ID, $K = 3$ EH receivers, $E_0 = 5$mW, $\gamma_0 = 10$dB

Fig. 7.8 presents the required transmit power $P_{TX}$ and total power consumption at the BS $P_{BS}$ of each technique with respect to the increasing SINR target of the ID, where the harvested energy requirement for each EH receiver is $E_0 = 5$mW. In Fig. 7.8 (a), it is observed that the FD case requires the lowest transmit power. For both of the proposed hybrid precoding, the partially-connected structures require a higher transmit power than fully-connected case, as only part of the channel can be exploited for the analog precoder. It is also observed that the proposed iterative hybrid design achieves an improved performance compared to the hybrid precoding based on SVD, especially for the partially-connected structures, and the iterative design with fully-connected structures can achieve a very close performance to the FD case due to the optimal analog precoding design. When we consider the total power consumption at the BS, on the contrary, it is interesting



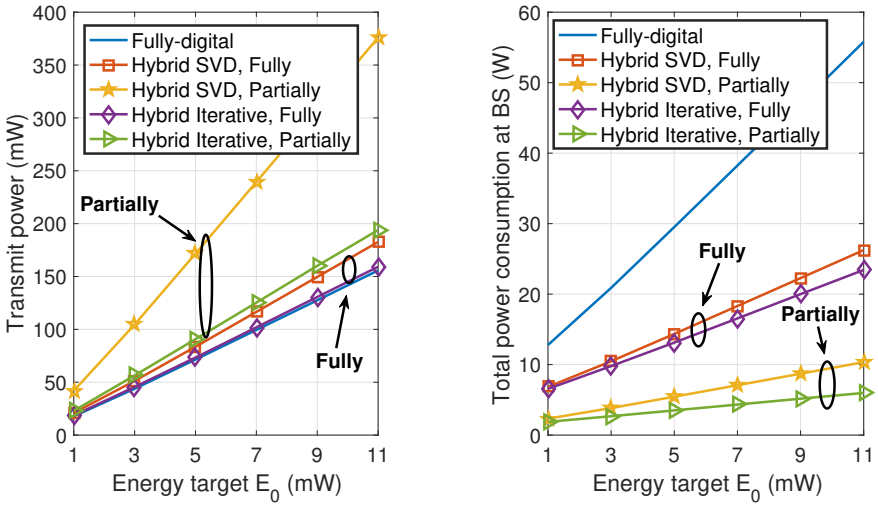

**(a)** Transmit power $P_{TX}$ (mW) **(b)** Total power at BS $P_{BS}$ (W)

**Figure 7.8:** Transmit power $P_{TX}$ and total power at the BS $P_{BS}$ required for $N_t = 12$, one ID, $K = 3$ EH receivers, $E_0 = 5$ mW, $N_{max} = 4$

to observe from Fig. 7.8 (b) that the FD case consumes the highest power, and the hybrid AD structures require a much lower power due to the reduced number of RF chains and phase shifters. The partially-connected structures require the lowest total power consumption at the BS and are therefore the most promising architectures for power efficient transmission.

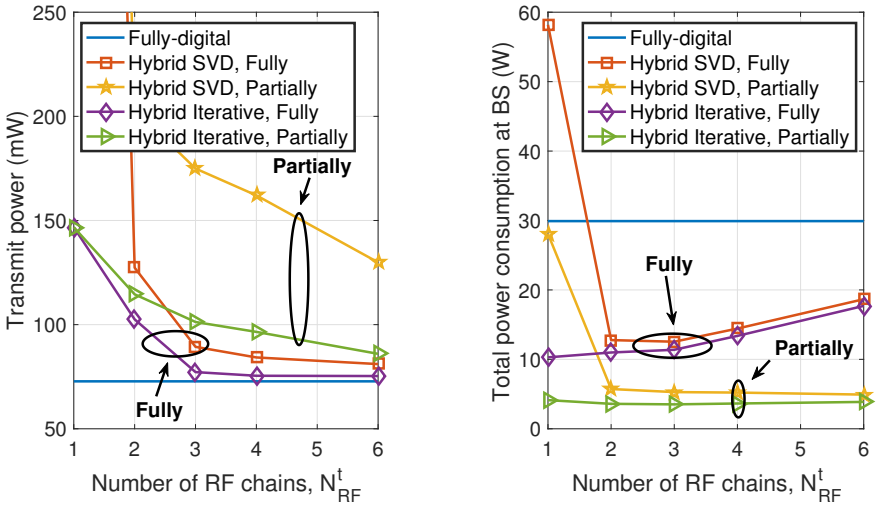
In Fig. 7.9, we compare the required transmit power and the total power at the BS with an increasing harvested energy requirement for each EH receiver, where the SINR target for the ID is $\gamma_0 = 10$dB. It can be observed that both the transmit power and the total power consumption at the BS are increasing with the increase in the harvested energy requirement $E_0$. In both figures, the proposed iterative algorithm outperforms the hybrid precoding based on SVD, and the performance gain is more significant for the partially-connected structures due to the optimal analog precoding design. In Fig. 7.9 (b), it is shown that the iterative hybrid method with partially-connected structures requires the lowest total power consumption at the BS.

In Fig. 7.10, we compare the performance of the hybrid approaches with respect to the number of RF chains. With a reduced number of RF chains,




**(a)** Transmit power $P_{TX}$ (mW) **(b)** Total power at BS $P_{BS}$ (W)

**Figure 7.9:** Transmit power $P_{TX}$ and total power at the BS $P_{BS}$ required for $N_t = 12$, one ID, $K = 3$ EH receivers, $\gamma_0 = 10$ dB, $N_{max} = 4$




**(a)** Transmit power $P_{TX}$ (mW) **(b)** Total power at BS $P_{BS}$ (W)

**Figure 7.10:** Transmit power $P_{TX}$ and total power at the BS $P_{BS}$ required for $N_t = 12$, one ID, $K = 3$ EH receivers, $\gamma_0 = 10$ dB, $N_{max} = 4$

the performance gap between the FD case and hybrid structures is larger, and the proposed iterative hybrid approach is shown to be less sensitive to the reduction in the number of RF chains. It is also observed that the performance gains of the proposed iterative algorithm over the low-complexity hybrid approach are more significant with a smaller number of RF chains.

In Fig. 7.10 (b), we note that the total power consumption at the BS is jointly decided by the transmit power in Fig. 7.10 (a) and the number of RF chains, where the proposed iterative algorithm still outperforms the low-complexity hybrid precoding based on SVD.

## 7.3 Summary

In this chapter, we first study the hybrid AD precoding techniques for mmWave communication systems, where the precoding methods are proposed by selecting the virtual paths. Three distinct algorithms are proposed, and each achieves a different performance-complexity tradeoff. Numerical results have shown that the proposed techniques with fully-connected structures can approach the FD precoding method. We further extend our study on the hybrid structure to energy-efficient MIMO SWIPT systems with limited RF chains, where we consider the scenario of one ID and several separate EH receivers. By analytically proving that only an information precoder is required in the scenario under study, we propose an iterative hybrid algorithm that exploits this observation. It is shown that the hybrid structures require much less power to achieve the same performance as the FD case. Both studies in this chapter have revealed that the hybrid structures are promising for the future multiple-antenna systems.

# Chapter 8

# Massive MIMO 1-Bit DAC Transmission with Symbol Scaling

This chapter is based on our works in [J1], [C1], [C2].

## 8.1 Introduction

In this chapter, we focus on the downlink precoding designs for the large-scale MIMO systems in the presence of 1-bit DACs. Towards 5G-and-beyond wireless communication systems, massive MIMO systems [9] have been shown to greatly improve the spectral efficiency, and therefore have received increasing research attention in recent years [9, 10]. Nevertheless, with a large number of antennas employed at the BS, the hardware complexity and the resulting power consumption have also increased dramatically, which hinders the practical implementation of massive MIMO. To achieve a compromise between performance, hardware complexity and the consequent power consumption, one potential technique for massive MIMO is to reduce the number of RF chains at the BS by employing the hybrid AD structure [11, 23], which has been introduced in Chapter 7.

In addition to the hybrid structures, another potential technique, which is the focus of this chapter, is to reduce the cost and power consumption per RF chain by employing very low-resolution DACs instead of high-precision DACs. It has been shown in [182] that DACs are one of the dominant power-

consuming hardware components in the downlink, whose power consumption grows exponentially with the resolution and linearly with the bandwidth. Therefore, employing low-resolution DACs, especially 1-bit DACs, can greatly reduce the power consumption per RF chain and the resulting total power consumed at the BS. When 1-bit DACs are employed, the output signal at each antenna element is equivalent to the constant-envelope symbol from a QPSK constellation, which enables the use of low-cost power amplifiers (PAs) and can further reduce the hardware complexity.

In the existing literature, most recent studies have focused on the performance analyses for massive MIMO uplink with low-resolution ADCs, especially for the 1-bit case [25, 26, 183], where it is shown that the number of quantisation bits can be reduced while a comparable performance is still achievable. For the case of downlink transmission with 1-bit DACs, there have been an increasing number of studies, due to the benefits mentioned above [16, 27, 28, 184, 185]. In [16], a simple quantised ZF method is considered, where the transmit signal vector is obtained by a direct quantisation on the ZF-precoded signals. The authors further analyse the performance of the quantised ZF method, and show that it outperforms the maximum likelihood (ML) encoder in the low-to-medium SNR regime. In [27, 184], the quantised linear precoding techniques based on MMSE are proposed, whose performance is shown to be superior to the quantised ZF approach in [16]. In [185], a non-linear symbol perturbation technique is introduced in 1-bit massive MIMO downlink for QPSK modulation, while in [28] an iterative non-linear precoding scheme named 'Pokemon' is introduced via a biconvex relaxation approach, where the proposed algorithm directly designs the transmit signal vector based on the MMSE criterion. Nevertheless, these MMSE-based approaches may be sub-optimal, as they ignore the fact that interference can be exploited on an instantaneous basis in [38, 41, 50, 51, 55].

In this chapter, we focus on the massive MIMO downlink systems and

consider multi-user precoding schemes with 1-bit DACs. For massive MIMO systems with 1-bit DACs, symbol-level operations are required as the output signals are dependent on the data symbols, which creates the opportunity to exploit the formulation of CI. It is this aspect of the transmission that allows us to observe the interference from an instantaneous point of view, and exploit it constructively [41, 50, 51]. We firstly propose a non-linear mapping scheme based on the CI formulation, where we directly design the quantised transmit signal vector. Nevertheless, due to the constraint on the output signals of 1-bit DACs, the resulting optimisation problem is shown to be non-convex. To solve this problem, we first apply a relaxation on the mathematical constraint resulting from the use of 1-bit DACs, such that the optimisation problem becomes convex. Subsequently, we apply an element-wise normalisation on the signal vector obtained from the relaxed optimisation to meet the constraint on the output signals of 1-bit DACs.

We further propose a low-complexity symbol scaling algorithm based on a coordinate transformation on the problem formulation of CI, where we directly select the output signal of 1-bit DACs for each antenna element on a sequential basis, and a relaxation is therefore no longer needed. The proposed symbol scaling approach consists of three stages: an initialisation stage where we decide the output signals for some antenna elements whose channel coefficients satisfy certain requirements, an allocation stage where we sequentially select the output signals for the residual antenna elements, and a refinement stage where we check whether the performance with the obtained signal vector can be further improved based on a greedy algorithm. Both the 'Sum-Max' and the 'Max-Min' criteria are considered in the allocation stage, and the output signal vector that returns the best performance is then obtained within the above two criteria. We further study the computational costs of the proposed optimisation-based and symbol scaling schemes in terms of the required floating-point operations. Numerical results reveal the advantages of the proposed techniques in terms of BER performance in

various scenarios. In terms of the computational complexity, the complexity of the symbol scaling scheme is negligible compared to that of the non-linear mapping approach, while its performance is superior to 'Pokemon' when their computational costs are similar, which favours its usefulness in practice.

For reasons of clarity, we summarise the contributions of this chapter as:

- We propose a non-linear mapping scheme for massive MIMO with 1-bit DACs, where we directly optimise the transmit signal vector. A two-step relaxation-normalisation is employed to solve the non-convex optimisation problem.
- Based on a coordinate transformation of the CI formulation, we further propose a low-complexity symbol scaling approach where we directly select the quantised signal on each antenna element via a three-stage process.
- We further study the computational costs of the optimisation-based non-linear mapping method and the symbol-scaling approach in terms of the required floating-point operations, which reveals the advantages of the proposed symbol-scaling approach.

## 8.2 System Model

We consider a massive MIMO downlink, where 1-bit DACs are employed at the BS, as depicted in Fig. 8.1. As we focus on the transmit-side process-

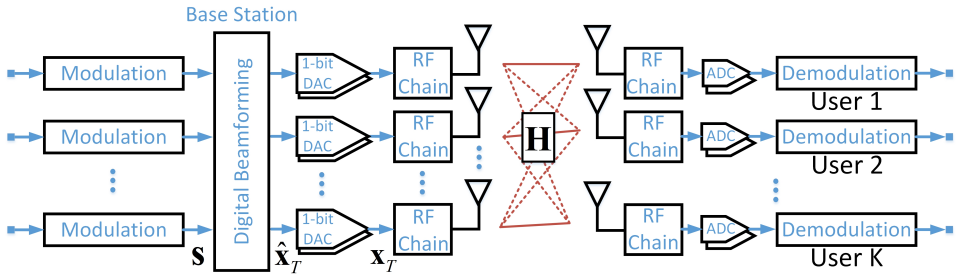



**Figure 8.1:** Massive MIMO downlink system model with 1-bit DACs

ing, ideal ADCs with infinite precision are assumed to be employed at each receiver. The BS with $N_t$ transmit antennas is communicating with $K$ single-antenna users simultaneously in the same time-frequency resource, where $K \ll N_t$. We focus on the downlink precoding designs and perfect CSI is assumed, while we also numerically study the performance of the proposed schemes with imperfect CSI in Chapter 8.6. Following the closely-related literature [16, 27, 184, 185], the symbol vector is assumed to be from a normalised PSK constellation. We denote the data symbol vector as $\mathbf{s} \in \mathbb{C}^{K\times 1}$, and the unquantised signal vector that is formed based on $\mathbf{s}$ as $\hat{\mathbf{x}}_T \in \mathbb{C}^{N_t\times 1}$. Then, the unquantised signal vector $\hat{\mathbf{x}}_T$ can be expressed as

$$\hat{\mathbf{x}}_T = \mathscr{B}(\mathbf{s}), \tag{8.1}$$

where $\mathscr{B}$ denotes a general transformation. When a linear precoding approach is employed, $\mathscr{B}$ represents the linear precoding matrix that is multiplied to $\mathbf{s}$ before quantisation, while in the case of non-linear precoding, $\mathscr{B}$ refers to a non-linear mapping scheme to form the transmit signals dependent on $\mathbf{s}$. With 1-bit DACs employed, the output signal vector is obtained as

$$\mathbf{x}_T = \mathscr{Q}(\hat{\mathbf{x}}_T). \tag{8.2}$$

In (8.2), $\mathscr{Q}$ denotes the 1-bit quantisation on both the real and imaginary part of each entry in $\hat{\mathbf{x}}_T$. We denote $x_n$ as the $n$-th entry in $\mathbf{x}_T$, and in this paper each $x_n$ is normalised to satisfy

$$x_n \in \left\{ \pm\frac{1}{\sqrt{2N_t}} \pm \frac{1}{\sqrt{2N_t}} \cdot j \right\}, \forall n \in \{1, 2, \cdots, N_t\}. \tag{8.3}$$

The above normalisation guarantees that $\|\mathbf{x}_T\|_F^2 = 1$, and accordingly we can express the received signal at the $k$-th user as

$$y_k = \sqrt{P} \cdot \mathbf{h}_k \mathbf{x}_T + n_k, \tag{8.4}$$

where $\mathbf{h}_k \in \mathbb{C}^{1\times N_t}$ denotes the flat-fading Rayleigh channel with each entry following a standard complex Gaussian distribution. $n_k$ denotes the additive Gaussian distributed noise at the receiver with zero mean and variance $\sigma^2$. $P$ is the total available transmit power for the antenna array, and for simplicity in this paper we assume uniform power allocation for the antenna array.

## 8.3 1-Bit Transmission Scheme based on Constructive Interference

In this section, we present the proposed optimisation-based methods by exploiting the CI. While the channel for massive MIMO is near-orthogonal, the imperfect hardware components in the form of 1-bit DACs introduce interference, which can be exploited by CI. To be more specific, we directly design the transmit signal vector $\mathbf{x}_T$ based on the CI formulation, which leads to the following optimisation

$$
\begin{aligned}
\mathscr{P}_{8.1}: \ & \max_{\mathbf{x}_T} \ t \\
s.t. \quad & \mathbf{h}_k \mathbf{x}_T = \lambda_k s_k, \ \forall k \in \mathscr{K} \\
& \left[\Re\left(\lambda_k\right) - t\right] \tan\theta_t \geq \left|\Im\left(\lambda_k\right)\right|, \ \forall k \in \mathscr{K} \\
& x_n \in \left\{ \pm\frac{1}{\sqrt{2N_t}} \pm \frac{1}{\sqrt{2N_t}} j \right\}, \ \forall n \in \{1, 2, \cdots, N_t\} \\
& t \geq 0
\end{aligned}
\tag{8.5}
$$

It is observed that the optimisation problem $\mathscr{P}_{8.1}$ is non-convex due to the output signal constraint for the 1-bit DACs in (8.5). To solve the above non-convex optimisation, we adopt a two-step approach.

### 8.3.1 Relaxation

In the first step, we relax the strict modulus constraint on each $x_n$ for both the real and imaginary part, and the resulting relaxed constraint can be ex-

pressed as

$$|\Re(x_n)| \leq \frac{1}{\sqrt{2N_t}},\ |\Im(x_n)| \leq \frac{1}{\sqrt{2N_t}},\ \forall n \in \{1,2,\cdots,N_t\}. \tag{8.6}$$

The optimisation problem $\mathscr{P}_{8.1}$ is reformulated into a relaxed version $\mathscr{P}_{8.2}$, given by

$$\begin{aligned}
\mathscr{P}_{8.2}:\ & \max_{\hat{\mathbf{x}}_T}\ t \\
s.t.\quad & \mathbf{h}_k\hat{\mathbf{x}}_T = \lambda_k s_k,\ \forall k \in \mathscr{K} \\
& [\Re(\lambda_k) - t]\tan\theta_t \geq |\Im(\lambda_k)|,\ \forall k \in \mathscr{K} \\
& |\Re(\hat{x}_n)| \leq \frac{1}{\sqrt{2N_t}},\ \forall n \in \{1,2,\cdots,N_t\} \\
& |\Im(\hat{x}_n)| \leq \frac{1}{\sqrt{2N_t}},\ \forall n \in \{1,2,\cdots,N_t\} \\
& t \geq 0
\end{aligned} \tag{8.7}$$

where we denote $\hat{x}_n$ as the $n$-th entry in the relaxed transmit signal vector $\hat{\mathbf{x}}_T$. The resulting $\mathscr{P}_{8.2}$ is convex and can be solved with convex optimisation tools.

### 8.3.2 Normalisation

The solution obtained from the relaxed optimisation $\mathscr{P}_{8.2}$ cannot always guarantee the equality on both the real and imaginary part of $\hat{x}_n$. To force the constraint of 1-bit transmission, the elements of the 1-bit DAC output $\mathbf{x}_T$ are obtained as

$$x_n = \frac{\Re(\hat{x}_n)}{\sqrt{2N_t}\cdot|\Re(\hat{x}_n)|} + \frac{\Im(\hat{x}_n)}{\sqrt{2N_t}\cdot|\Im(\hat{x}_n)|}\cdot j,\ \forall n \in \{1,2,\cdots,N_t\}. \tag{8.8}$$

We further note that, while we perform a relaxation on the 1-bit DAC constraint on each $x_n$ in $\mathscr{P}_{8.1}$, it turns out that most entries of the obtained $\hat{\mathbf{x}}_T$ from the relaxed problem $\mathscr{P}_{8.2}$ already meet the requirement for 1-bit quantisation, i.e. strict equality is satisfied for these entries in the case of massive MIMO. Accordingly, there exist only a small number of $\hat{x}_n$ that need to be normalised to satisfy the constraint on the output signals for 1-bit DACs.

| Antenna number $N_t$ | 16 | 32 | 48 | 64 | 80 | 96 | 112 | 128 |
|---|---|---|---|---|---|---|---|---|
| Ratio $\eta$ | 20.52% | 10.8% | 7.28% | 5.46% | 4.37% | 3.65% | 3.13% | 2.73% |

**Table 8.1:** $\eta$ with respect to the number of transmit antennas, $K = 4$, 500 channel realisations

Moreover, to evaluate the deviation of the relaxed optimisation $\mathscr{P}_{8.2}$ from the original problem $\mathscr{P}_{8.1}$, we define $n_{\Re}$ and $n_{\Im}$ as the number of entries in the obtained $\hat{\mathbf{x}}_T$, whose absolute values are smaller than $\frac{1}{\sqrt{2N_t}}$ for the real and imaginary part, respectively. We further introduce

$$\eta = \frac{n_{\Re} + n_{\Im}}{2N_t} \tag{8.9}$$

as the ratio of the number of entries that do not satisfy the 1-bit transmission to the total number of entries in $\hat{\mathbf{x}}_T$, and this ratio therefore represents the deviation of the solution obtained by the relaxed problem from the original problem. It is obtained that $0 \leq \eta \leq 1$, and $\mathscr{P}_{8.2}$ is equivalent to $\mathscr{P}_{8.1}$ if $\eta = 0$. It is also observed that a smaller value of $\eta$ means that the relaxed optimisation is closer to the original optimisation.

To study this numerically, we present the value of $\eta$ with respect to the number of antennas in Table 8.1, where we have assumed a total number of $K = 4$ users in the downlink system, and the result is based on 500 channel realisations. It is observed that the ratio $\eta$ decreases with the increase in the number of transmit antennas, which means that the solution obtained via the relaxed optimisation problem $\mathscr{P}_{8.2}$ can be regarded as asymptotically optimal with an increasing number of transmit antennas in the case of massive MIMO.

## 8.4 Proposed Low-Complexity Symbol Scaling Approach

While the above non-linear mapping method can be relaxed into a convex optimisation problem, the corresponding computational complexity is still

prohibitively high as the variable dimension is equal to the number of transmit antennas. We study this mathematically and numerically in Chapter 8.5 and 8.6, respectively. Therefore in this section, we propose a three-stage symbol scaling approach, whose computational cost is much lower. It will be shown in the numerical results that for the small-scale MIMO systems, the low-complexity method even outperforms the optimisation-based nonlinear mapping scheme and achieves the best performance, since no relaxation or normalisation is required for this approach. In the case of massive MIMO, it also achieves a comparable performance to the proposed nonlinear mapping scheme.

### 8.4.1 A New Look at the Constructive Interference Criteria

To introduce the proposed symbol scaling scheme, we first perform a coordinate transformation on the constructive interference constraint. To be specific, for $\mathbb{N}$-PSK modulations, each data symbol in the conventional real-imaginary plane can be expressed as

$$s_{(l)} = e^{j\cdot\left[\frac{2\pi}{\mathbb{N}}(l-1)+\frac{\pi}{4}\right]},\ l \in \{1,2,\cdots\mathbb{N}\}, \tag{8.10}$$

where $s_{(l)}$ denotes the $l$-th constellation point. Given the constellation points, the equations that represent the two detection thresholds for a specific con-

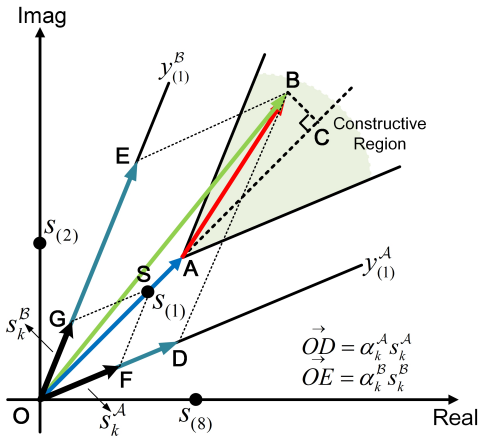



**Figure 8.2:** Decomposition along the detection thresholds for 8PSK

stellation point $s_{(l)}$ can be expressed as

$$\begin{aligned} y_{(l)}^{\mathbb{A}} &= \tan\left[\frac{2\pi}{\mathbb{N}}(l-1)+\frac{\pi}{4}-\frac{\pi}{\mathbb{N}}\right]\cdot x = \tan\left[\frac{2\pi}{\mathbb{N}}\cdot l+\frac{\pi}{4}-\frac{3\pi}{\mathbb{N}}\right]\cdot x, \\ y_{(l)}^{\mathbb{B}} &= \tan\left[\frac{2\pi}{\mathbb{N}}(l-1)+\frac{\pi}{4}+\frac{\pi}{\mathbb{N}}\right]\cdot x = \tan\left[\frac{2\pi}{\mathbb{N}}\cdot l+\frac{\pi}{4}-\frac{\pi}{\mathbb{N}}\right]\cdot x. \end{aligned} \tag{8.11}$$

For the proposed symbol-scaling methods, without loss of generality we assume the data symbol for user $k$ is $s_k = s_{(l)}$. We then propose to decompose the constellation points along their corresponding two detection thresholds, expressed as

$$s_k = s_{(l)} = s_k^{\mathbb{A}} + s_k^{\mathbb{B}}, \tag{8.12}$$

where $s_k^{\mathbb{A}}$ is parallel to $y_{(l)}^{\mathbb{A}}$ and $s_k^{\mathbb{B}}$ is parallel to $y_{(l)}^{\mathbb{B}}$. Accordingly, $s_k^{\mathbb{A}}$ and $s_k^{\mathbb{B}}$ can be expressed as

$$s_k^{\mathbb{A}} = \frac{e^{j\cdot\left(\frac{2\pi}{\mathbb{N}}\cdot l+\frac{\pi}{4}-\frac{3\pi}{\mathbb{N}}\right)}}{\rho} = A_k^{\Re} + j\cdot A_k^{\Im},\ \ s_k^{\mathbb{B}} = \frac{e^{j\cdot\left(\frac{2\pi}{\mathbb{N}}\cdot l+\frac{\pi}{4}-\frac{\pi}{\mathbb{N}}\right)}}{\rho} = B_k^{\Re} + j\cdot B_k^{\Im}, \tag{8.13}$$

where $\left(A_k^{\Re}, A_k^{\Im}\right)$ and $\left(B_k^{\Re}, B_k^{\Im}\right)$ denote the coordinates of the bases $s_k^{\mathbb{A}}$ and $s_k^{\mathbb{B}}$ in the real-imaginary plane, respectively. The constant $\rho$ is a scaling factor to guarantee that $s_k = s_k^{\mathbb{A}} + s_k^{\mathbb{B}}$. Note that for a normalized $\mathbb{N}$-PSK modulation, $|s_k| = 1$ and $\rho$ is accordingly obtained as

$$\rho = \left| e^{j\cdot\left(\frac{2\pi}{\mathbb{N}}\cdot l+\frac{\pi}{4}-\frac{3\pi}{\mathbb{N}}\right)} + e^{j\cdot\left(\frac{2\pi}{\mathbb{N}}\cdot l+\frac{\pi}{4}-\frac{\pi}{\mathbb{N}}\right)} \right|. \tag{8.14}$$

The above decomposition is also shown geometrically in 8.2, where we employ a 8-PSK modulation as an example. Specifically, for the considered constellation point in 8.2, we obtain $\vec{OS} = s_{(1)}$, which further leads to

$$s_k^{\mathbb{A}} = \frac{e^{j\cdot\left(\frac{2\pi}{8}\cdot 1+\frac{\pi}{4}-\frac{3\pi}{8}\right)}}{\rho} = \frac{e^{j\cdot\frac{\pi}{8}}}{\left| e^{j\cdot\frac{\pi}{8}} + e^{j\cdot\frac{3\pi}{8}} \right|},\ s_k^{\mathbb{B}} = \frac{e^{j\cdot\left(\frac{2\pi}{8}\cdot 1+\frac{\pi}{4}-\frac{\pi}{8}\right)}}{\rho} = \frac{e^{j\cdot\frac{3\pi}{8}}}{\left| e^{j\cdot\frac{\pi}{8}} + e^{j\cdot\frac{3\pi}{8}} \right|}. \tag{8.15}$$

Then for each $k$, instead of employing a complex scaling value $\lambda_k$ that is multiplied by $s_k$, with the above formulation (8.12)-(8.14) we introduce a

symbol scaling approach where we decompose the noiseless received signal for user $k$ along the two corresponding detection thresholds of $s_k$, given by

$$\mathbf{h}_k \mathbf{x}_T = \alpha_k^{\mathbb{A}} s_k^{\mathbb{A}} + \alpha_k^{\mathbb{B}} s_k^{\mathbb{B}}, \tag{8.16}$$

where

$$\alpha_k^{\mathbb{A}} \geq 0,\ \alpha_k^{\mathbb{B}} \geq 0,\ \forall k \in \mathscr{K} \tag{8.17}$$

are scaling factors. We observe that a larger value of $\alpha_k^{\mathbb{A}}$ or $\alpha_k^{\mathbb{B}}$ represents a larger distance to the detection threshold, and by expanding (45) using the coordinate transformation, we can obtain a generic expression of $\alpha_k^{\mathbb{A}}$ and $\alpha_k^{\mathbb{B}}$ as a function of the transmit signal vector, given by (see Appendix D)

$$\begin{aligned}
\alpha_k^{\mathbb{A}} &= \frac{B_k^{\Im}\mathbf{h}_k^{\Re} - B_k^{\Re}\mathbf{h}_k^{\Im}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}}\mathbf{x}_T^{\Re} - \frac{B_k^{\Im}\mathbf{h}_k^{\Im} + B_k^{\Re}\mathbf{h}_k^{\Re}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}}\mathbf{x}_T^{\Im}, \\
\alpha_k^{\mathbb{B}} &= \frac{A_k^{\Re}\mathbf{h}_k^{\Im} - A_k^{\Im}\mathbf{h}_k^{\Re}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}}\mathbf{x}_T^{\Re} + \frac{A_k^{\Re}\mathbf{h}_k^{\Re} + A_k^{\Im}\mathbf{h}_k^{\Im}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}}\mathbf{x}_T^{\Im}.
\end{aligned} \tag{8.18}$$

In (8.18), for simplicity we have employed the following notation

$$\mathbf{x}_T^{\Re} = \Re\left(\mathbf{x}_T\right),\ \mathbf{x}_T^{\Im} = \Im\left(\mathbf{x}_T\right),\ \mathbf{h}_k^{\Re} = \Re\left(\mathbf{h}_k\right),\ \mathbf{h}_k^{\Im} = \Im\left(\mathbf{h}_k\right). \tag{8.19}$$

By further denoting

$$\begin{aligned}
\mathbf{A}_k &= \frac{B_k^{\Im}\mathbf{h}_k^{\Re} - B_k^{\Re}\mathbf{h}_k^{\Im}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}},\ \mathbf{B}_k = -\frac{B_k^{\Im}\mathbf{h}_k^{\Im} + B_k^{\Re}\mathbf{h}_k^{\Re}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}}, \\
\mathbf{C}_k &= \frac{A_k^{\Re}\mathbf{h}_k^{\Im} - A_k^{\Im}\mathbf{h}_k^{\Re}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}},\ \mathbf{D}_k = \frac{A_k^{\Re}\mathbf{h}_k^{\Re} + A_k^{\Im}\mathbf{h}_k^{\Im}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}},
\end{aligned} \tag{8.20}$$

the formulation of (8.18) is simplified to

$$\alpha_k^{\mathbb{A}} = \mathbf{A}_k\mathbf{x}_T^{\Re} + \mathbf{B}_k\mathbf{x}_T^{\Im},\ \alpha_k^{\mathbb{B}} = \mathbf{C}_k\mathbf{x}_T^{\Re} + \mathbf{D}_k\mathbf{x}_T^{\Im}. \tag{8.21}$$

By defining

$$\mathbf{R}_k = \begin{bmatrix} \mathbf{A}_k & \mathbf{B}_k \end{bmatrix},\ \mathbf{I}_k = \begin{bmatrix} \mathbf{C}_k & \mathbf{D}_k \end{bmatrix}, \tag{8.22}$$

and

$$\mathbf{x} = \left[ \begin{array}{cc} \left(\mathbf{x}_T^{\Re}\right)^T & \left(\mathbf{x}_T^{\Im}\right)^T \end{array} \right]^T, \ \mathbf{\Lambda} = \left[\alpha_1^{\mathbb{A}}, \cdots, \alpha_K^{\mathbb{A}}, \alpha_1^{\mathbb{B}}, \cdots, \alpha_K^{\mathbb{B}}\right]^T, \tag{8.23}$$

(8.18) can be further expressed in a compact form as

$$\mathbf{\Lambda} = \mathbf{M}\mathbf{x}, \tag{8.24}$$

where $\mathbf{M}$ is given by

$$\mathbf{M} = \left[ \begin{array}{cccccc} \mathbf{R}_1^T & \cdots & \mathbf{R}_K^T & \mathbf{I}_1^T & \cdots & \mathbf{I}_K^T \end{array} \right]^T. \tag{8.25}$$

With the above formulation, we can then construct the optimization problem as

$$\begin{aligned} \mathscr{P}_{8.3}: \ & \max_{\mathbf{x}} \min_{l} \ \alpha_l \\ s.t. \ \ & \mathbf{\Lambda} = \mathbf{M}\mathbf{x} \\ & \alpha_l \geq 0, \ \forall l \in \mathscr{L} \\ & x_i^E \in \left\{ \frac{1}{\sqrt{2N_t}}, -\frac{1}{\sqrt{2N_t}} \right\}, \ \forall i \in \mathscr{I} \end{aligned} \tag{8.26}$$

where we have omitted $\Re$ and $\Im$ in the expression of the entries of $\mathbf{\Lambda}$, and simply denote $\alpha_l$ as its $l$-th entry. In $\mathscr{P}_{8.3}$, $\mathscr{L} = \{1, 2, \cdots, 2K\}$, $x_i^E$ denotes the $i$-th entry in $\mathbf{x}$ and $\mathscr{I} = \{1, 2, \cdots, 2N_t\}$. The above optimization problem $\mathscr{P}_{8.3}$ is interpreted as follows: we aim to maximize the minimum value of $\alpha_l$ by selecting each $x_i^E$ as either $\frac{1}{\sqrt{2N_t}}$ or $-\frac{1}{\sqrt{2N_t}}$. With the above problem formulation, the relaxation-normalization process on the transmit signals is no longer needed. The above formulation motivates us to propose the following low-complexity scheme, which consists of three stages: an initialization stage, an allocation stage, and a refinement stage, each presented in detail below.

### 8.4.2 Initialisation Stage

In the initialisation stage, we directly select the value of $x_i^E$ for some $i$ by simple observation. To achieve this, we firstly rewrite (8.24) into

$$\mathbf{\Lambda} = \sum_{i=1}^{2N_t} \mathbf{M}_i x_i^E, \tag{8.27}$$

where we decompose $\mathbf{M}$ into

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 & \cdots & \mathbf{M}_{2N_t} \end{bmatrix}, \tag{8.28}$$

with each $\mathbf{M}_i \in \mathbb{C}^{2K\times 1}$. Based on (8.27), we have the following observation.

**Observation**: As long as all the entries of $\mathbf{M}_i$ share the same sign, then it is optimal to set the sign of the corresponding $x_i^E$ equal to that of $\mathbf{M}_i$, as in this case the values of each entry in $\mathbf{\Lambda}$ are guaranteed to increase.

Accordingly, the corresponding $x_i^E$ is obtained as

$$x_i^E = \frac{\text{sgn}\left(\mathbf{M}_i\right)}{\sqrt{2N_t}},\ \forall i \in \mathscr{U}, \tag{8.29}$$

where $\text{sgn}\left(\mathbf{a}\right)$ defines a vector sign function and is only valid when each entry in the vector $\mathbf{a}$ has the same sign. $\mathscr{U}$ denotes the set that consists of the column indices of $\mathbf{M}$ that satisfy the sign-identity condition. We further introduce a column vector $\mathbf{t}$ that represents a temporary value of $\mathbf{\Lambda}$, given by

$$\mathbf{t} = \sum_{i\in\mathscr{V}} \mathbf{M}_i x_i^E, \tag{8.30}$$

where the set $\mathscr{V}$ consists of the column indices of $\mathbf{M}$ whose corresponding $x_i^E$ have been allocated a value. We note that when $card\left(\mathscr{V}\right) = 2N_t$, we have $\mathbf{t} = \mathbf{\Lambda}$.

In the case that no column in $\mathbf{M}$ satisfies the sign-identity condition, in the initialisation stage we select only one column, i.e. $card\left(\mathscr{U}\right) = 1$, with the

following criterion:

$$i = \underset{i \in \mathscr{H}}{\arg\max} \left\| \mathbf{M}_i \right\|_1, \tag{8.31}$$

which selects the column that has the maximum effect on the value of $\mathbf{\Lambda}$. Then, the value of the corresponding $x_i^E$ is set as

$$x_i^E = \frac{\operatorname{sgn}\left( \left\| \mathbf{M}_i \right\|_1 \right)}{\sqrt{2N_t}}. \tag{8.32}$$

In the initialisation stage, we have $\mathscr{V} = \mathscr{U}$ or $card\left(\mathscr{V}\right) = 1$. We summarise the algorithm for the initialisation stage in Algorithm 8.1.

**Algorithm 8.1** Initialisation Stage

**input** : $\mathbf{s}$, $\mathbf{H}$
**output** : $\mathbf{t}$, $\mathscr{V}$
Decompose each $s_k = s_k^{\Re} + s_k^{\Im}$ based on modulation type;
Obtain $\mathbf{M}$ based on (8.12)-(8.25);
Find $\mathbf{M}_i$ that satisfies the sign-identity condition;
Obtain $\mathscr{U}$;
**if** $\mathscr{U} \neq \emptyset$ **then**
    $x_i^E = \frac{\operatorname{sgn}(\mathbf{M}_i)}{\sqrt{2N_t}}$, $\forall i \in \mathscr{U}$;
    $\mathscr{V} = \mathscr{U}$;
**else**
    Obtain $i$ based on (8.31), $x_i^E = \frac{\operatorname{sgn}(\|\mathbf{M}_i\|_1)}{\sqrt{2N_t}}$;
    $\mathscr{V} = \{i\}$;
**end if**
Calculate $\mathbf{t}$ based on (8.30).

### 8.4.3 Allocation Stage

At this stage, we allocate the value of each $x_i^E$ for the residual $i$ that belongs to $\mathscr{W}$, where we define the set $\mathscr{W}$ as

$$\mathscr{W} = \left\{ i \,|\, i \in \mathscr{H} \text{ and } i \notin \mathscr{V} \right\}. \tag{8.33}$$

$\mathscr{W}$ consists of those $x_i^E$ whose values have not been allocated in the initialisation stage. In the following allocation stage, we consider both a 'Sum-Max' and a 'Max-Min' criteria for the allocation process.

**Sum-Max:** For the allocation scheme based on the 'Sum-Max' criterion, instead of considering a max-min optimisation as in $\mathscr{P}_{8.3}$, we consider an alternative sum-max optimisation, where the objective function is constructed as

$$\mathscr{F}(\mathbf{x}) = \text{sum}(\mathbf{\Lambda}), \tag{8.34}$$

where $\text{sum}(\mathbf{a})$ returns the sum of the entries in a column vector $\mathbf{a}$. Based on (8.24), the objective can be further transformed into

$$\mathscr{F}(\mathbf{x}) = \mathbf{m}\mathbf{x} = \sum_{i=1}^{2N_t} [\mathbf{m}]_i x_i^E, \tag{8.35}$$

where $\mathbf{m} \in \mathbb{C}^{1\times 2N_t}$ is the sum of the entries in each row of $\mathbf{M}$. Each $[\mathbf{m}]_i$ denotes the $i$-th entry in $\mathbf{m}$, given by

$$[\mathbf{m}]_i = \sum_{l=1}^{2K} [\mathbf{M}_i]_l. \tag{8.36}$$

It is easy to observe that $\mathscr{F}(\mathbf{x})$ is maximised when the sign of each $x_i^E$ is the same as that of $[\mathbf{m}]_i$, and therefore the optimal $x_i^E$ for the 'Sum-Max' criterion is given by

$$x_i^E = \frac{\text{sgn}\{[\mathbf{m}]_i\}}{\sqrt{2N_t}}, \ \forall i \in \mathscr{W}. \tag{8.37}$$

The algorithm for the allocation stage based on 'Sum-Max' is summarised in Algorithm 8.2.

**Remark:** While the above solution guarantees that the sum of $\alpha_l$ is maximised, it does not specifically consider each value of $\alpha_l$, which may lead to a performance loss. Indeed, it is possible that the value of one $\alpha_l$ can be very small or even negative. This is the reason why the refinement in Chapter 8.4.4 is further introduced.

**Max-Min:** For the 'Max-Min' allocation criterion, in each step we aim to improve the minimum value in $\mathbf{\Lambda}$ as much as possible. Denoting $q$ as the row index of the minimum entry in $\mathbf{t}$ obtained in the initialisation stage, we

**Algorithm 8.2** Allocation Stage - 'Sum-Max'

**input** : $\mathscr{V}$, $\mathbf{M}$
**output** : $\mathbf{x}_{\text{sum}-\text{max}}$
Calculate $\mathscr{W}$ based on (8.33);
Calculate $\mathbf{m}$ and each $\mathbf{m}(i)$ based on (8.35), (8.36);
Allocate $x_i^E = \frac{\text{sgn}[\mathbf{m}(i)]}{\sqrt{2N_t}}$, $\forall i \in \mathscr{W}$;
Obtain $\mathbf{x}$, denoted as $\mathbf{x}_{\text{sum}-\text{max}}$.

obtain

$$[\mathbf{t}]_q = \min(\mathbf{t}), \tag{8.38}$$

where $\min(\mathbf{t})$ returns the minimum value in $\mathbf{t}$. Subsequently, we iteratively select $\mathbf{M}_i$ with the largest absolute value in the $q$-th row, given by

$$i = \underset{i \in \mathscr{W}}{\arg\max} \left| [\mathbf{M}_i]_q \right|, \tag{8.39}$$

and the corresponding $x_i^E$ is then obtained as

$$x_i^E = \frac{\text{sgn}[\mathbf{M}_i(q)]}{\sqrt{2N_t}}. \tag{8.40}$$

Then, we update $\mathscr{V}$ and $\mathbf{t}$, and based on the updated $\mathbf{t}$ we repeat the above procedure until $\mathscr{V} = \mathscr{H}$. This means that each entry in $\mathbf{x}$ has been allocated, and the algorithm for the allocation stage based on 'Max-Min' is summarised in Algorithm 8.3.

**Algorithm 8.3** Allocation Stage - 'Max-Min'

**input** : $\mathscr{V}$, $\mathbf{M}$, $\mathbf{t}$
**output** : $\mathbf{x}_{\text{max}-\text{min}}$
**while** $\mathscr{V} \neq \mathscr{H}$ **do**
    Calculate $\mathscr{W}$ based on (8.33);
    Obtain $q$ that satisfies $\mathbf{t}(q) = \min(\mathbf{t})$;
    Find $i = \underset{i \in \mathscr{W}}{\arg\max} \left| [\mathbf{M}_i]_q \right|$;
    Allocate $x_i^E = \frac{\text{sgn}[\mathbf{M}_i(q)]}{\sqrt{2N_t}}$;
    Update $\mathscr{V}$ and $\mathbf{t}$;
**end while**
Obtain $\mathbf{x}$, denoted as $\mathbf{x}_{\text{max}-\text{min}}$.

### 8.4.4 Refinement Stage

In the refinement stage, we check whether the performance based on the obtained signal vector in the allocation stage can be further improved based on a greedy algorithm. To introduce the refinement process, we denote the obtained expanded 1-bit signal vector after the allocation stage as $\mathbf{x}$ (obtained based on either the 'Sum-Max' or the 'Max-Min' criterion). First, we sequentially change the sign of one entry (for example $x_i^E$) in $\mathbf{x}$ at a time while fixing the signs of other entries in $\mathbf{x}$, and denote the modified signal vector as $\mathbf{x}_{(i)}$. We then compare the minimum value in $\mathbf{\Lambda}$ obtained by the modified $\mathbf{x}_{(i)}$ with the minimum value in the original $\mathbf{\Lambda}$ obtained by $\mathbf{x}_{(0)}$. The sign of $x_i^E$ is selected as the one that returns a larger minimum value in $\mathbf{\Lambda}$. The refinement process is sequentially performed for each entry in $\mathbf{x}_{(0)}$. The algorithm for the refinement stage is then shown in Algorithm 8.4.

**Algorithm 8.4** Refinement Stage

**input** : $\mathbf{x}_{\text{sum}-\text{max}}$ (or $\mathbf{x}_{\text{max}-\text{min}}$)
**output** : $\mathbf{x}_T$
Denote $\mathbf{x}_{(0)} = \mathbf{x}_{\text{sum}-\text{max}}$ (or $\mathbf{x}_{\text{max}-\text{min}}$);
**for** $i = 1 : 2N_t$ **do**
  Calculate $\mathbf{\Lambda}_{(0)} = \mathbf{M}\mathbf{x}_{(0)}$;
  Obtain $\mathbf{x}_{(i)} = \left[x_1^E, \ldots, x_{i-1}^E, -x_i^E, x_{i+1}^E, \ldots, x_{2N_t}^E\right]^T$;
  Calculate $\mathbf{\Lambda}_{(i)} = \mathbf{M}\mathbf{x}_{(i)}$;
  **if** $\min\left(\mathbf{\Lambda}_{(i)}\right) > \min\left(\mathbf{\Lambda}_{(0)}\right)$ **then**
    $x_i^E \leftarrow -x_i^E$;
    Update $\mathbf{x}_{(0)}$;
  **end if**
**end for**
Obtain $\mathbf{x}_T$ based on the updated $\mathbf{x}_{(0)}$.

The refinement stage is performed for the signal vectors obtained by both the 'Sum-Max' and 'Max-Min' criteria independently. Accordingly, the final output signal vector of the proposed symbol scaling approach that generates the best performance is selected between the signal vectors obtained with these two criteria.

### 8.4.5 Algorithm

Based on the above description, the algorithm for the three-stage symbol scaling scheme is summarised in Algorithm 8.5, where the final output signal vector of the proposed symbol scaling method that generates the best performance is selected within the signal vectors obtained by the 'Sum-Max' and 'Max-Min' criteria.

**Algorithm 8.5** The Proposed Symbol Scaling Approach

**input:** $\mathbf{s}$, $\mathbf{H}$
**output:** $\mathbf{x}_T$
**Initialisation Stage**
Obtain $\mathscr{V}$, $\mathbf{M}$, and $\mathbf{t}$ with Algorithm 8.1;
**Allocation Stage**
**1. 'Sum-Max':**
Obtain $\mathbf{x}_{\text{sum}-\text{max}}$ with Algorithm 8.2;
**2. 'Max-Min':**
Obtain $\mathbf{x}_{\text{max}-\text{min}}$ with Algorithm 8.3;
**Refinement Stage**
Update both $\mathbf{x}_{\text{sum}-\text{max}}$ and $\mathbf{x}_{\text{max}-\text{min}}$ with Algorithm 8.4;
Calculate $\mathbf{\Lambda}_{\text{s}} = \mathbf{M}\mathbf{x}_{\text{sum}-\text{max}}$ and $\mathbf{\Lambda}_{\text{m}} = \mathbf{M}\mathbf{x}_{\text{max}-\text{min}}$;
**if** $\min(\mathbf{\Lambda}_{\text{s}}) > \min(\mathbf{\Lambda}_{\text{m}})$ **then**
  $\mathbf{x} = \mathbf{x}_{\text{sum}-\text{max}}$;
**else**
  $\mathbf{x} = \mathbf{x}_{\text{max}-\text{min}}$;
**end if**
Decompose $\mathbf{x} = \left[ \begin{array}{cc} \left(\mathbf{x}_T^{\Re}\right)^T & \left(\mathbf{x}_T^{\Im}\right)^T \end{array} \right]^T$;
Output $\mathbf{x}_T = \mathbf{x}_T^{\Re} + \mathbf{x}_T^{\Im} \cdot j$.

## 8.5 Computational Complexity Analysis

In this section we study the computational costs of the proposed schemes in terms of the required number of real multiplications and additions. As a reference, we also study the complexity of the exhaustive search scheme and the non-linear 'Pokemon' approach in [28]. For the optimization-based approach, the complexity is evaluated based on the number of arithmetic operations [188].

### 8.5.1 Exhaustive Search

For massive MIMO transmission with 1-bit quantization, the output signal on each antenna element has 4 potential values, i.e., each $x_n \in \left\{\frac{1}{\sqrt{2}}+j\cdot\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}-j\cdot\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}+j\cdot\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}-j\cdot\frac{1}{\sqrt{2}}\right\}$. The exhaustive search method first searches all the possible signal combinations and then selects the best one, which means that there are a total number of $4^{N_t}$ signal combinations. For each signal combination, it takes $4KN_t$ real multiplications and $2K\left(2N_t-1\right)$ real additions to compute $\mathbf{\Lambda}$ based on (8.24), as $\mathbf{M}\in\mathbb{C}^{2K\times 2N_t}$. Therefore, the total number of required operations for the exhaustive search scheme considering all the possible combinations is obtained as

$$\mathrm{C_E} = \left(4KN_t + 4KN_t - 2K\right)\cdot 4^{N_t} = \left(8KN_t - 2K\right)\cdot 2^{2N_t}. \tag{8.41}$$

It is easy to conclude that in the case of massive MIMO, the exhaustive search scheme is inapplicable due to the overwhelmingly high computational cost.

### 8.5.2 Symbol Scaling Approach

In the following we calculate the computational cost for each stage of the proposed symbol scaling approach. For both allocation criteria, the main computational cost in the initialization and allocation stage comes from the calculation of $\mathbf{t}\in\mathbb{C}^{2K\times 1}$ based on (8.30). While the calculation of $\mathbf{t}$ is not necessary for the 'Sum-Max' criterion, we note that $\mathbf{t}$ is required in the refinement stage. Each additional $\left(\mathbf{m}_i x_i^E\right)$ term that is added to $\mathbf{t}$ requires $2K$ multiplications and $2K$ additions, and $\mathbf{t}$ is updated $2N_t$ times after the allocation stage, where we note $\mathbf{M}\in\mathbb{C}^{2K\times 2N_t}$. The resulting computational cost is

$$\mathrm{C_L^1} = 2N_t\left(2K+2K\right) = 8KN_t. \tag{8.42}$$

Moreover, for the 'Max-Min' allocation criterion, we need to iteratively allocate the value for the residual $x_i^E$, which introduces an additional computational cost for 'Max-Min' in the allocation stage. Since $card\left(\mathscr{V}\right)$ is difficult to obtain analytically in the initialization stage, we consider a worst-case com-

plexity where $card(\mathscr{V}) = 1$, and in each iteration obtaining $q$ and $i$ in Algorithm 8.3 requires $2K$ and $2N_t$ operations respectively, which leads to

$$\mathrm{C}_{\mathrm{L}}^{2} = (2N_t - 1)(2K + 2N_t) \approx 4N_t^2 + 4KN_t \tag{8.43}$$

in the case of massive MIMO. In the refinement stage, it is easy to observe that the initial $\mathbf{\Lambda}_{(0)} = \mathbf{t}$. Then, in each iteration of Algorithm 8.4 we only need to calculate the corresponding $\mathbf{m}_i \cdot \left(-x_i^E\right)$ and include it in $\mathbf{\Lambda}_{(i)}$. For each $x_i^E$ this takes $2K$ multiplications and $2K$ additions, and therefore the computational cost for the refinement stage is

$$\mathrm{C}_{\mathrm{L}}^{3} = 2N_t(2K + 2K) = 8KN_t. \tag{8.44}$$

Based on Algorithm 8.5, both $\mathbf{x}_{\mathrm{sum-max}}$ and $\mathbf{x}_{\mathrm{max-min}}$ should be refined. Accordingly, we can obtain the total computational cost for the proposed symbol scaling approach as

$$\begin{aligned} \mathrm{C}_{\mathrm{L}} &= \mathrm{C}_{\mathrm{L}}^{1} + \mathrm{C}_{\mathrm{L}}^{2} + 2\mathrm{C}_{\mathrm{L}}^{3} \\ &= 8KN_t + 4N_t^2 + 4KN_t + 2 \times 8KN_t \\ &= 4N_t^2 + 28KN_t. \end{aligned} \tag{8.45}$$

### 8.5.3 Pokemon

As a comparison, we also include the complexity of the non-linear 'Pokemon' scheme proposed in [28]. The 'Pokemon' approach is based on biconvex relaxation, whose performance is dependent on the number of required iterations. Based on [28], in each iteration we need to first calculate a vector $\mathbf{z} \in \mathbb{C}^{2N_t \times 1}$ based on $\mathbf{z} = \mathbf{U}\mathbf{x}$ where $\mathbf{U} \in \mathbb{C}^{2N_t \times 2N_t}$, and then update the signal vector $\mathbf{x} \in \mathbb{C}^{2N_t \times 1}$ with a projection function. The calculation of $\mathbf{z}$ requires a total of $4N_t^2$ multiplications and $2N_t(2N_t - 1)$ additions, while the update of $\mathbf{x}$ requires $4N_t$ multiplications. Assuming a maximum number of iterations

$n_{\max}$, this leads to

$$\begin{aligned} \mathrm{C_P} &= n_{\max}\left(4N_t^2 + 4N_t^2 - 2N_t + 4N_t\right) \\ &= n_{\max}\left(8N_t^2 + 2N_t\right). \end{aligned} \tag{8.46}$$

Comparing the computational cost of 'Pokemon' with the proposed symbol scaling method, we have

$$\frac{\mathrm{C_L}}{\mathrm{C_P}} = \frac{4N_t^2 + 28KN_t}{n_{\max}\left(8N_t^2 + 2N_t\right)} = \frac{2N_t + 14K}{n_{\max}\left(4N_t + 1\right)}. \tag{8.47}$$

In the case of massive MIMO where $K$ is finite while the antenna number goes to infinity, (8.47) simplifies to approximately

$$\frac{\mathrm{C_L}}{\mathrm{C_P}} = \frac{2 + \frac{14K}{N_t}}{n_{\max}\left(4 + \frac{1}{N_t}\right)} \approx \frac{1}{2n_{\max}}. \tag{8.48}$$

With the fact $n_{\max} \geq 1$, generally we obtain $\mathrm{C_L} < \mathrm{C_P}$.

### 8.5.4 Optimisation-based Non-linear Mapping $\mathscr{P}_{8.1}$

For the proposed non-linear mapping scheme, it is difficult to calculate the required number of multiplications and additions. Therefore, we resort to [188] and evaluate its complexity based on the arithmetic complexity.

For this non-convex optimization problem, the complexity is dominated by solving the relaxed convex problem $\mathscr{P}_{8.2}$ via the interior-point method [124]. Based on our reformulated $\mathscr{P}_{8.3}$ in Chapter 8.4.1, we first express the equivalent real representation of $\mathscr{P}_{8.2}$ in a standard form as

$$\begin{aligned} \mathscr{P}_{8.4}: \ & \max_{\mathbf{v}} \ \mathbf{c}^T\mathbf{v} \\ s.t. \ & \mathbf{q}_l\mathbf{v} \leq 0, \forall l \in \mathscr{L} \\ & -\mathbf{e}_{i+1}^T\mathbf{v} \leq \frac{1}{\sqrt{2N_t}}, \ -\mathbf{e}_{i+1}^T\mathbf{v} \leq \frac{1}{\sqrt{2N_t}}, \ \forall i \in \mathscr{I} \\ & \mathbf{v} = \left[t, x_1^E, x_2^E, \cdots, x_{2N_t}^E\right]^T \\ & \mathbf{c} = [1, 0, 0, \cdots, 0]^T \end{aligned} \tag{8.49}$$

In $\mathscr{P}_{8.4}$, $\mathbf{q}_l = \begin{bmatrix} 1 & -\hat{\mathbf{m}}_l \end{bmatrix}$, where $\hat{\mathbf{m}}_l$ denotes the $l$-th row of $\mathbf{M}$. Based on [188], the arithmetic complexity bound of the above optimization via the interior-point method is given by

$$\mathrm{C_N} = (M_c + N_v)^{1.5} N_v^2 \cdot \mathrm{D}(\mathbf{p}, \varepsilon), \tag{8.50}$$

where $\varepsilon$ is the accuracy of the solution, $N_v$ denotes the dimension of the variable $\mathbf{v}$, and $M_c$ is the total number of the constraints in the optimization. Based on the construction of $\mathscr{P}_{8.4}$, we obtain

$$M_c = 4N_t + 2K, \;\; N_v = 2N_t + 1, \tag{8.51}$$

which further leads to the expression of $\mathrm{C_N}$ as

$$\mathrm{C_N} = (6N_t + 2K + 1)^{1.5} (2N_t + 1)^2 \cdot \mathrm{D}(\mathbf{p}, \varepsilon). \tag{8.52}$$

$\mathrm{D}(\mathbf{p}, \varepsilon)$ is the number of digits of accuracy for a solution with the accuracy $\varepsilon$, and is given by

$$\mathrm{D}(\mathbf{p}, \varepsilon) = \ln\left(\frac{\mathrm{Dim}(\mathbf{p}) + \|\mathbf{p}\|_1 + \varepsilon^2}{\varepsilon}\right), \tag{8.53}$$

where the column vector $\mathbf{p}$ represents a permutation vector that contains the parameters in both the objective function and the constraints [188]. For our considered problem $\mathscr{P}_{8.4}$, $\mathbf{p}$ is given as

$$\mathbf{p} = \Bigg[ (2K+4N_t), (2N_t+1), 1, \underbrace{1, \cdots, 1}_{2K}, \underbrace{(-\hat{\mathbf{m}}_1), (-\hat{\mathbf{m}}_2), \cdots, (-\hat{\mathbf{m}}_{2K})}_{2K}, \underbrace{1, \cdots, 1}_{2N_t}, \underbrace{-1, \cdots, -1}_{2N_t}, \underbrace{\frac{1}{\sqrt{2N_t}}, \cdots, \frac{1}{\sqrt{2N_t}}}_{4N_t} \Bigg]^T, \tag{8.54}$$

which further leads to

$$\|\mathbf{p}\|_1 = \sqrt{10N_t + 4K + 2\sqrt{2N_t} + \|\mathbf{M}\|_F^2 + 2}. \tag{8.84}$$

In (8.53), $\mathrm{Dim}(\mathbf{p})$ denotes the dimension of the permutation vector $\mathbf{p}$, and is accordingly obtained as

$$\begin{aligned}\mathrm{Dim}(\mathbf{p}) &= (M_c+1)(N_v+2)+2 \\ &= (4N_t+2K+1)(2N_t+2)+2.\end{aligned} \tag{8.85}$$

Given the expressions for $\mathrm{Dim}(\mathbf{p})$ and $\|\mathbf{p}\|_1$, we arrive at the final expression of the complexity for $\mathscr{P}_{8.2}$, given by

$$\mathrm{C_N} = (6N_t+2K+1)^{1.5}(2N_t+1)^2 \cdot \ln\left(\frac{(4N_t+2K+1)(2N_t+2)+\sqrt{10N_t+4K+2\sqrt{2N_t}+\|\mathbf{M}\|_F^2+2}+\varepsilon^2}{\varepsilon}\right). \tag{8.86}$$

## 8.6 Numerical Results

In this section we present the numerical results of the proposed approaches based on Monte Carlo simulations. In each plot, the transmit SNR is defined as $\gamma = P/\sigma^2$. Both QPSK and 8-PSK modulations are considered in the numerical results. We compare our proposed methods with both the quantised linear approaches and the non-linear mapping algorithms, and for clarity the following abbreviations are used throughout this section:

- 'ZF-FD': Unquantised ZF precoding with infinite-precision DACs;
- 'ZF 1-Bit': Quantised ZF approach with 1-bit DACs in [16];
- 'MMSE': MMSE-based quantised linear scheme in [27];
- 'Pokemon, $n_{\max} = K$': Non-linear Pokemon algorithm proposed in [28] with $K$ iterations;
- 'Constructive': Proposed non-linear mapping scheme $\mathscr{P}_{8.1}$;
- 'sum-max': Proposed symbol scaling approach based on the 'sum-max' allocation scheme with Algorithm 8.1, 8.2 and 8.4;
- 'max-min': Proposed symbol scaling approach based on the 'max-min' allocation method with Algorithm Algorithm 8.1, 8.3 and 8.4;
- 'Symbol Scaling': Proposed symbol scaling method obtained via Algorithm 8.5, where we select the best signal vector out of 'sum-max' and 'max-min' criteria.

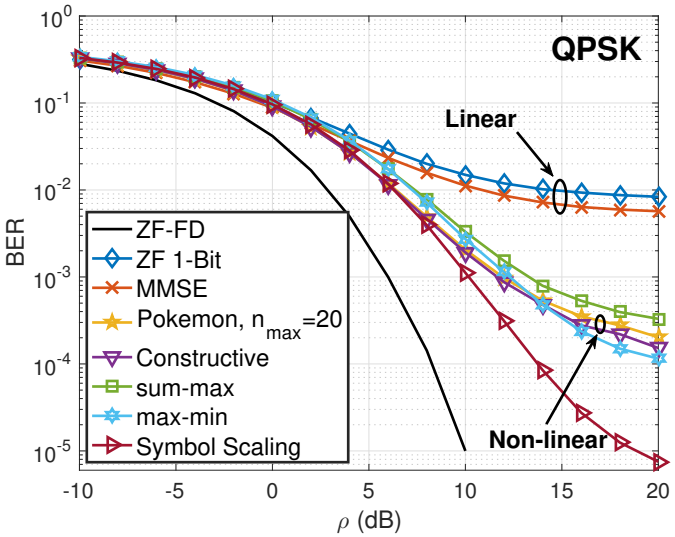



**Figure 8.3:** BER v.s. transmit SNR, $N_t = 8$, $K = 2$, $n_{\max} = 20$, QPSK

In Fig. 8.3, we firstly consider a moderate-scale MIMO with a total number of $N_t = 8$ transmit antennas at the BS and $K = 2$ single-antenna users in the system. For approaches with 1-bit quantisation, we observe that the proposed symbol scaling method based on Algorithm 8.5 achieves the best BER performance, while both the proposed non-linear mapping scheme and 'Pokemon' achieve an inferior performance. This is because both the non-linear mapping method and the 'Pokemon' approach involve the relaxation-normalisation process. For small-scale MIMO systems, based on Table 8.1 we can infer that $\eta$ will be large in this case, which means that the deviation of the solution obtained by the relaxation-normalisation process from the solution of the original 1-bit optimisation problem is large, and the normalisation process may lead to further detection errors. For the proposed symbol scaling approach, the performance is promising since we directly select the quantised signal for each antenna element and therefore no relaxation or quantisation is needed.

We proceed to consider a massive MIMO system with $N_t = 128$ transmit antennas and $K = 16$ users in Fig. 8.4. In the case of massive MIMO, all the schemes can achieve a lower BER thanks to the large number of antennas at the BS, and generally non-linear schemes outperform linear schemes.

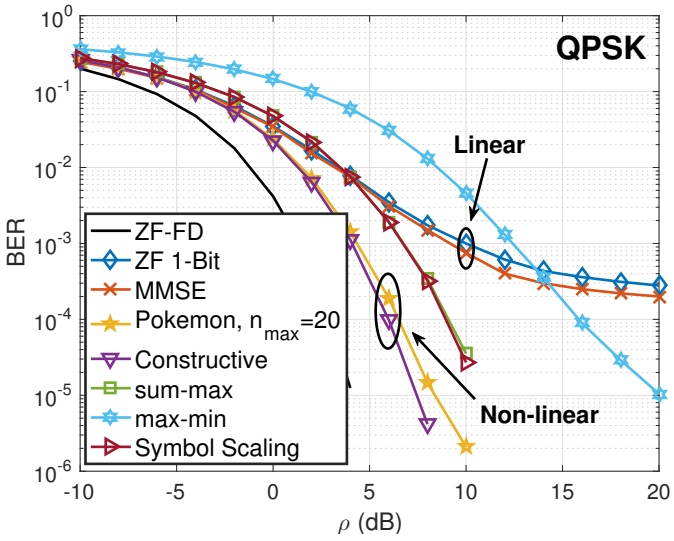



**Figure 8.4:** BER v.s. transmit SNR, $N_t = 128$, $K = 16$, $n_{\max} = 20$, QPSK

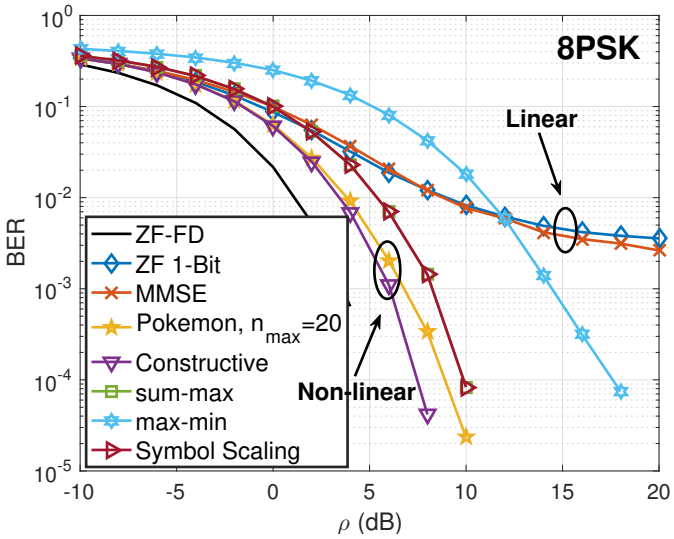



**Figure 8.5:** BER v.s. transmit SNR, $N_t = 128$, $K = 8$, $n_{\max} = 20$, 8-PSK

For approaches with 1-bit DACs, the proposed non-linear mapping method outperforms the non-linear 'Pokemon' algorithm and achieves the best BER performance. As for the proposed low-complexity symbol scaling approach, by comparing Fig. 8.3 and Fig. 8.4, we can observe that the 'Max-Min' criterion is most suitable for small-scale MIMO systems, while the 'Sum-Max' criterion is more favourable for massive MIMO systems.

In Fig. 8.5, we show the performance of different schemes for 8-PSK

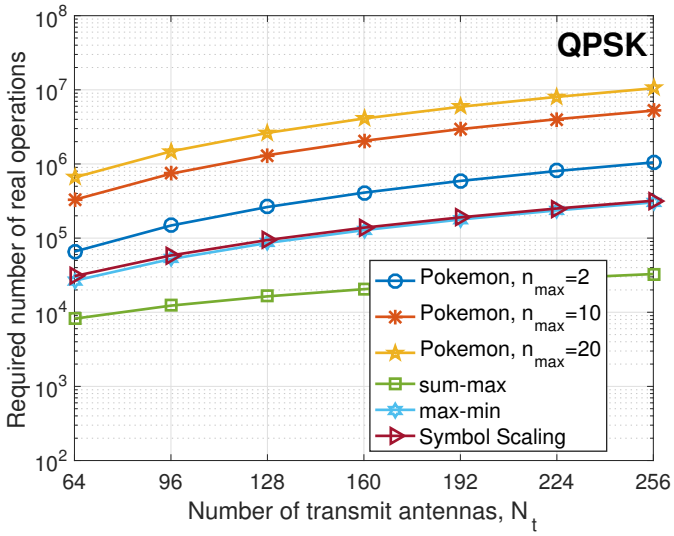



**Figure 8.6:** Analytical computational cost comparison, $K = 8$, $n_{\max} = 20$, QPSK

modulation with $N_t = 128$ and $K = 8$. For 1-bit quantised precoding approaches, it is observed that the proposed optimisation-based non-linear scheme achieves the best BER performance. For the symbol scaling approach, it is observed that in the case of 8-PSK, only a 1dB SNR loss is observed compared to the non-linear iterative 'Pokemon' algorithm, and therefore the proposed low-complexity symbol scaling approach is more favourable in terms of the performance and complexity tradeoff.

In Fig. 8.6, we compare the computational complexity of each approach in terms of the required number of real multiplications and additions. It is observed that the computational cost of the proposed symbol-scaling method based on sum-max requires the lowest computational cost, while the number of operations required for the proposed symbol scaling approach is much smaller than the required number of operations for 'Pokemon', even for the case of $n_{\max} = 2$ iterations. The complexity gains of the proposed symbol scaling approach therefore favour its practical application.

To further compare the proposed schemes with 'Pokemon', in Fig. 8.7 we present the BER performance with different number of iterations for Pokemon. The number of iterations does not have an effect on other methods and therefore the BER for the other methods remains constant. It is observed

that the performance of Pokemon improves as $n_{\max}$ increases. Nevertheless, we note that the improvement becomes less significant with a larger $n_{\max}$ and Pokemon achieves its best performance when $n_{\max}$ is around 25. An important observation is when $n_{\max} = 2, 3$, where the computational cost of Pokemon and our proposed scheme is similar, as shown by (**??**), and our proposed symbol scaling approach is shown to achieve an improved performance, which validates the superiority of the proposed approach.

To demonstrate the performance-complexity tradeoff directly, in Fig. 8.8 we depict the BER with respect to the required number of real operations for a range of transmit antennas from $N_t = 32$ to $N_t = 128$, where the number of users is fixed as $K = 8$. We observe a significant gain for the 'Symbol Scaling' approach compared to 'Pokeman' in terms of the performance-complexity tradeoff, especially when the number of antennas is large. Moreover, while the performance of the proposed low-complexity method based on 'sum-max' achieves an inferior performance to the 'Symbol Scaling' approach when $N_t$ is large, it indeed achieves a better BER performance with a lower computational cost compared to Pokemon with $n_{\max} = 2$. Both of the above observations indicate an advantage for the proposed scheme based on

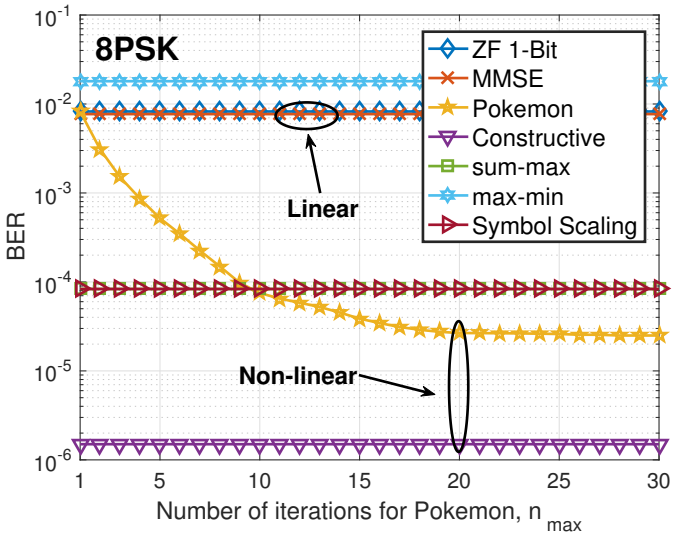



**Figure 8.7:** BER v.s. Pokemon iteration number $n_{\max}$, $N_t = 128$, $K = 8$, $\rho = 10$dB, 8-PSK

symbol scaling.

All the above results are based on the assumption of perfect CSI. In the following, we numerically investigate the performance of the proposed approaches with imperfect CSI. The channel estimation techniques for massive MIMO with 1-bit quantisation is an ongoing topic of research [183, 186], and an exact model for the imperfect CSI for this scenario is still not known.

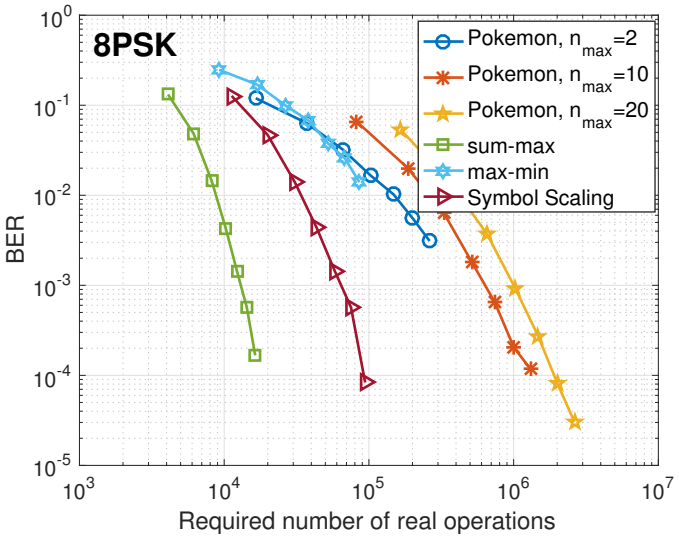



**Figure 8.8:** BER v.s. analytical computational costs, $K = 8$, $\rho = 10$dB, 8-PSK

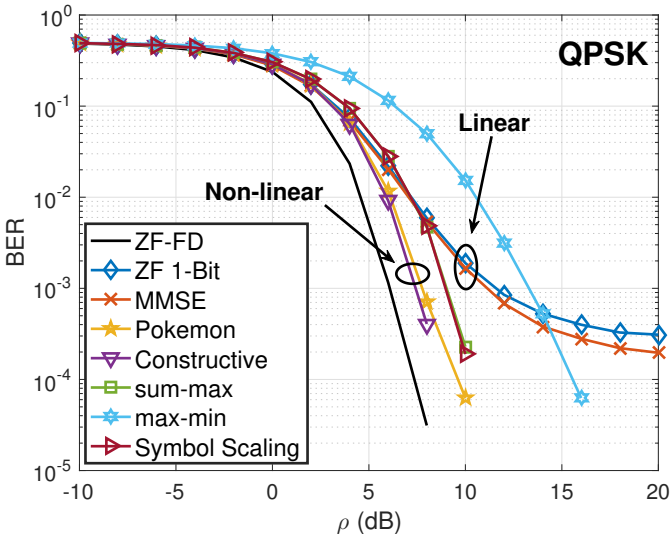



**Figure 8.9:** BER v.s. transmit SNR, $N_t = 128$, $K = 16$, $n_{\max} = 20$, QPSK, Imperfect CSI, $\beta = 2.5$

Therefore, in the following we employ a generic CSI model for TDD systems, where the BS only has knowledge of a noisy version of $\mathbf{H}$, given by

$$\hat{\mathbf{H}} = \mathbf{H} + \mathbf{Q}. \tag{8.87}$$

In (8.87), $\hat{\mathbf{H}}$ is the obtained CSI at the BS. $\mathbf{Q}$ denotes an error matrix with $\mathbf{Q} \sim \mathbb{CN}(\mathbf{0}, \delta \cdot \mathbf{I})$, where $\delta$ denotes the variance of the channel error. $\delta$ is modelled as inversely proportional to the transmit SNR and is expressed as $\delta = \beta/\rho$, where $\beta$ denotes the error coefficient [38]. The BER result with imperfect CSI is depicted in Fig. 8.9, where a similar trend can be observed. We can further observe that the proposed non-linear mapping method still achieves the best performance among the schemes with 1-bit quantisation in the case of imperfect CSI, while the proposed low-complexity symbol scaling approach can achieve a comparable performance with a greatly reduced computational cost.

## 8.7 Summary

In this chapter, we propose the downlink precoding approaches for the massive MIMO downlink with 1-bit DACs based on the formulation of CI, and we consider both a quantised linear scheme and a non-linear mapping scheme. With the analyses of Lagrangian and KKT conditions, the linear method is mathematically proven to be equivalent to the quantised ZF precoding. For the proposed non-linear mapping approach, it is shown to be non-convex and solved by a two-step relaxation-quantisation process. We further propose a low-complexity symbol scaling approach, where the quantised transmit signals are directly obtained. Numerical results reveal the superiority of the proposed symbol scaling approach in small-scale MIMO systems. In the case of massive MIMO, the performance advantage of the proposed non-linear mapping method is validated, while the proposed symbol scaling approach achieves a better performance-complexity tradeoff, which favours its usefulness in practical systems.

# Chapter 9

# Conclusions and Future Work

Downlink precoding and transmit beamforming techniques are essential to achieve spatial multiplexing for multiple-antenna systems. Meanwhile, taking into consideration the realistic hardware components and their corresponding effects on the designs of downlink transmission methods is also of great necessity for the future 5G-and-beyond wireless communication systems. Accordingly, this thesis has studied and proposed a number of hardware-informed transmission strategies for realistic BS structures with compact antenna arrays, reduced number of RF chains, and low-precision DACs.

## 9.1 Conclusions

In this thesis, Chapter 2 presents an overview of multiple-antenna communication systems. Conventional downlink transmission approaches are described, including closed-form precoding schemes and optimisation-based transmit beamforming methods. After identifying the main drawbacks of the above techniques and highlighting areas that need further investigations, the main contributions of this thesis are presented in Chapter 3 - Chapter 8. To be more specific,

- Chapter 3 has proposed two novel VP precoding techniques for the multiple-modulation scenarios. It is shown that the proposed techniques achieve significant performance gains over the existing BDVP

and UGVP designed for multiple-modulation cases. The main observations in this chapter can be summarised as:

C3.1 The significant performance gain of the conventional VP technique over linear precoding methods is extended to multiple-modulation scenarios by the proposed JVP precoding. With a simple but effective constellation scaling approach, the JVP method achieves a similar performance to single-modulation VP scheme without incurring significant additional computational costs.

C3.2 Thanks to the CI formulation, the introduced JCVP precoding can achieve a similar performance to the JVP precoding in multiple-modulation cases, while requires a reduced computational cost, especially for the users employing PSK modulations. The complexity gains make the proposed JCVP precoding more attractive for multiple-modulation scenarios, compared to the JVP precoding.

- Both Chapter 4 and Chapter 5 have studied the exploitation of MC effect by the use of tunable antenna loads in conventional compact multiple-antenna systems, where Chapter 4 focuses on P2P MIMO systems by the exploiting the CI formulation, while Chapter 5 extends to multi-user scenarios by the proposition of a joint AD beamforming method. It is shown that the performance of compact MIMO systems can be further improved with the employment of tunable antenna loads. The central remarks within these two chapters are:

C4.1 For a P2P MIMO system considered in Chapter 4, by employing a ZF equalisation at the receiver, the MC matrix can be regarded as an interference matrix at the transmitter. It is further proven that full elimination of the coupling effect is not achievable by the proposed approach.

C4.2 The CI formulation is particularly useful in optimising the tunable antenna loads, which not only returns a feasible solution with practical antenna values, thanks to the constructive region that expands the feasible region of the tunable loads, but also achieves a significant SNR gain in the BER performance.

C5.1 For a multi-user case in Chapter 5, both the proposed joint and decoupled approach return an improved BER performance, even compared to the ideal case assuming no coupling effect. Specifically, the decoupled approach further enables the exploitation of MC with closed-form precoders. The above observations reveal the benefits of exploiting the coupling effect instead of compensating for it in compact multiple-antenna systems.

C5.2 The proposed sequential search algorithm is promising when realistic quantised load values are considered, which achieves a close-to-optimal performance while only requires a very low computational cost. It is particularly effective when the number of transmit antennas is large or the quantisation level is high, where the exhaustive search method is too complicated for practical implementation.

- In Chapter 6, by taking realistic quantised tunable loads into consideration, we design the quantisation-robust precoding techniques specifically for the compact parasitic-antenna arrays. It is proven that the optimality is achieved by optimising the feeding voltages only, where convex optimisation approaches are employed to obtain the optimal feeding voltage in the presence of quantised antenna loads. The main observations obtained from the results of this chapter are as follows:

C6.1 The introduction of quantisation in the tunable load values has a significant impact on the performance of compact parasitic antenna arrays. An error floor is observed in medium-to-high SNR

regimes. Specifically, even a small quantisation interval can lead to a significant performance loss.

C6.2 The proposed quantisation-robust method can well alleviate the performance degradation by quantisation, especially for multi-fed parasitic arrays. With a larger quantisation interval, the performance gain of the proposed technique over conventional cases is more significant. Particularly, the derived closed-form expression for single-fed parasitic arrays motivates the efficient implementation of the proposed technique.

- Chapter 7 first focuses on the hybrid precoding in large-scale mmWave communication systems, and propose three distinct hybrid precoding designs based on the virtual path selection. This chapter further extends the study on the hybrid precoding to energy-efficient MIMO SWIPT systems, where both a low-complexity and an iterative methods are proposed. The central conclusions drawn from this chapter are:

C7.1 Among all of the three proposed methods for hybrid precoding in mmWave communications, the joint design achieves the best performance with the highest computational cost. Compared to the joint design, only a slight performance loss is observed for the sub-optimal method, which requires the least complexity and may be the most promising approach in practical mmWave systems.

C7.2 The advantage of employing hybrid structures extends to the small-scale MIMO SWIPT systems in terms of power efficiency. Both of the proposed hybrid designs achieve an improved performance over the FD case, while further gains are observed for the proposed iterative algorithm over the low-complexity method based on SVD. In particular, the required total power consumption at the BS reveals that the partially-connected hybrid struc-

tures are more advantageous over the fully-connected hybrid structures.

- In Chapter 8, we study the downlink precoding for large-scale multiple-antenna systems with low-precision DACs based on the CI formulation. The quantised CI-based linear precoding method is mathematically shown to be equivalent to the quantised ZF method for massive MIMO, while the proposed non-linear method can significantly improve the performance. A low-complexity symbol-scaling method that directly selects the quantised transmit signals is further proposed. The results observed in this chapter lead to the following remarks:

  C8.1 Both the quantised linear and non-linear approaches include a relaxation-normalisation process, while the symbol-scaling approach directly designs the quantised transmit signals. Therefore in small-scale MIMO systems, the proposed method based on symbol scaling achieves the best BER performance, as it is very likely that the normalisation process leads to a detection error when the number of transmit antennas is small.

  C8.2 While the performance is inferior to the non-linear method in the case of massive MIMO, the proposed symbol scaling approach achieves an improved performance and complexity tradeoff, which makes it more appealing in practical large-scale multiple-antenna systems with 1-bit DACs.

## 9.2 Future Work

The proposed techniques in this thesis have motivated further investigations in some research directions, especially in the area of future large-scale multiple-antenna systems. In particular, the following research lines are of interest to the author for the future work:

- **Downlink transmission design for BSs with a reduced number of RF chains and 1-bit DACs:** Compared to conventional FD systems, the results and observations in both Chapter 7 and Chapter 8 have demonstrated the potential gains in employing hardware-efficient BSs. Current research has either focused on the hybrid structures with high-precision DACs or considered low-precision DACs with full RF chains, both of which can achieve a higher energy efficiency performance than the FD systems. As a further step, it is therefore of particular interest to propose novel transmission methods where the joint use of limited number of RF chains and low-resolution DACs is considered. To be more specific, while the downlink precoding designs for 1-bit DACs do not need to consider the power scaling factor, this factor is needed when the hybrid structure is further included, which requires further investigations and may motivate new transmission strategies.

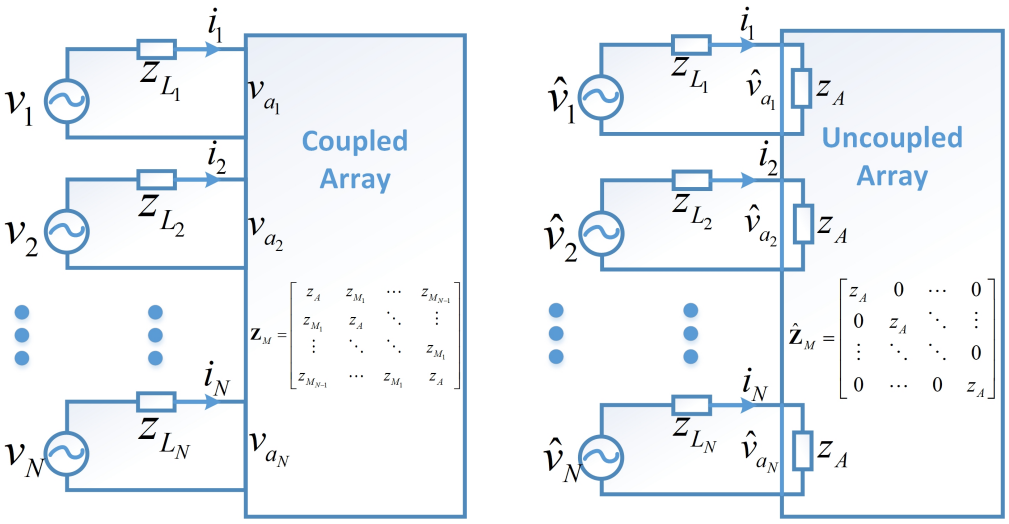
- **Energy efficiency maximisation for hybrid BS structures with low-precision DACs:** Achieving energy-efficient transmissions is one important task for the future wireless communication systems, due to the increasing power consumption in the wireless BSs and devices. While both the use of small number of RF chains and low-resolution DACs can increase the energy efficiency, maximising the energy efficiency is not the primal target. Therefore, it is appealing to directly consider the maximisation on the energy efficiency. To achieve this, in the first step an analytical result on the system capacity needs to be derived in the presence of limited RF chains and finite-precision DACs, which may not be easy and requires deep investigations. An optimisation on the required number of RF chains and DAC precision to maximise the energy efficiency can be further considered, and the tradeoff between the number of RF chains and the quantisation bits in the DACs is also an interesting topic.

- **Hybrid beamforming designs for SWIPT systems with power-splitting receivers:** In Chapter 7 of this thesis, we have considered the application of hybrid precoding in MIMO SWIPT systems. Nevertheless, a simple SWIPT architecture with separate ID and the EH receivers is considered. As a future work, it is particularly interesting to study the combination of hybrid precoding with power-splitting receivers, where each user can simultaneously decode data symbols and harvest energy. More specifically, it will be interesting to consider a joint optimisation on the hybrid precoders and the power splitting ratio to achieve a specific target, for example, power minimisation, SINR balancing, harvested energy maximisation or energy efficiency maximisation. This study may further motivate the use of hybrid structures in small-scale multiple-antenna systems due to their more favourable power, hardware costs, and performance tradeoff.

To finally summarise, this thesis has presented several hardware-informed downlink transmission approaches, each designed for a specific transmitter structure. The author hopes that the results and observations obtained within this thesis can help motivate novel designs, which respect the realistic hardware components in the multiple-antenna systems for the future 5G-and-beyond wireless communication systems.

# Appendices

## Appendix A: Derivation of the Mutual Coupling Matrix with Tunable Loads




**(a)** Coupled Array **(b)** Equivalent Uncoupled Array

**Figure A.1:** Circuit representation of a coupled array and the uncoupled transformation

The derivation of the mutual coupling matrix with tunable loads is based on [64]. An array consisting of $N$ coupled antennas is considered where we assume each antenna element is connected to a RF chain and a voltage source. The representation of the array and its equivalent uncoupled transformation is shown in Fig. A.1. Accordingly, based on the Ohm's

law we obtain

$$\mathbf{v} = \mathbf{v}_L + \mathbf{v}_a = diag\,(\mathbf{z}_L)\,\mathbf{i} + \mathbf{Z}_M\mathbf{i} = [diag\,(\mathbf{z}_L) + \mathbf{Z}_M]\,\mathbf{i}, \tag{A.1}$$

where $\mathbf{v} = [v_1, v_2, \cdots, v_N]$ denotes the source voltage vector that is considered constant, $\mathbf{v}_L = [v_{L_1}, v_{L_2}, \cdots, v_{L_N}]$ is the load voltage vector, $\mathbf{v}_a = [v_{a_1}, v_{a_2}, \cdots, v_{a_N}]$ is the voltage vector on the antenna load, $\mathbf{z}_L = [z_{L_1}, z_{L_2}, \cdots, z_{L_N}]$ denotes the load impedance vector, and $\mathbf{Z}_M$ is the mutual impedance matrix, given by

$$\mathbf{Z}_M = \begin{bmatrix} z_{1,1} & z_{1,2} & \cdots & z_{1,N} \\ z_{2,1} & z_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & z_{N-1,N} \\ z_{N,1} & \cdots & z_{N,N-1} & z_{N,N} \end{bmatrix}. \tag{A.2}$$

Without loss of generality, following [64, 187] we further assume that the self-impedance of the antenna is independent of the presence of other antenna element and all antennas are identical, which leads to

$$z_{1,1} = z_{2,2} = \cdots = z_{N,N} = z_A, \tag{A.3}$$

and a fixed mutual impedance value for two arbitrary antenna elements with an identical antenna spacing. Accordingly, the mutual impedance matrix is further simplified into

$$\mathbf{Z}_M = \begin{bmatrix} z_A & z_{M_1} & \cdots & z_{M_{N-1}} \\ z_{M_1} & z_A & \ddots & \vdots \\ \vdots & \ddots & \ddots & z_{M_1} \\ z_{M_{N-1}} & \cdots & z_{M_1} & z_A \end{bmatrix}. \tag{A.4}$$

Based on , (A.1) can be further transformed into

$$\mathbf{i} = [diag\,(\mathbf{z}_L) + \mathbf{Z}_M]\,\mathbf{v}^{-1}. \tag{A.5}$$

In the equivalent transformation, $\hat{\mathbf{v}}$ is denoted as the equivalent source voltage vector to obtain the same radiation fields as the coupled model. The mutual coupling matrix can be defined as

$$\hat{\mathbf{v}} = \mathbf{Z}\mathbf{v}. \tag{A.6}$$

Based on the uncoupled transformation, the equivalent voltage for each antenna impedance is obtained in matrix form as

$$\hat{\mathbf{v}}_a = z_A \cdot \mathbf{i} \tag{A.7}$$

with the same currents as in the coupled model. By substituting (A.5) into (A.7), (A.7) is further transformed into

$$\frac{1}{z_A} \cdot \hat{\mathbf{v}}_a = \left[ diag\left(\mathbf{z}_L\right) + \mathbf{Z}_M \right]^{-1} \mathbf{v}. \tag{A.8}$$

With

$$\hat{\mathbf{v}}_a = diag\left( \left[ \frac{z_A}{z_{L_1} + z_A}, \cdots, \frac{z_A}{z_{L_N} + z_A} \right] \right) \hat{\mathbf{v}}, \tag{A.9}$$

we further obtain

$$\hat{\mathbf{v}} = \underbrace{\left[ z_A \cdot \mathbf{I} + diag\left(\mathbf{z}_L\right) \right] \left[ \mathbf{Z}_M + diag\left(\mathbf{z}_L\right) \right]^{-1}}_{\mathbf{Z}} \mathbf{v} \tag{A.10}$$

which defines the mutual coupling matrix $\mathbf{Z}$.

# Appendix B: Calculation of the Antenna Load $z_A$ and Mutual Impedance Matrix $\mathbf{Z}_M$

Based on [63], the antenna impedance $z_A$ and the mutual impedance $z_{M_k}$ can be calculated with EMF methods based on the antenna spacing $d$. For the antenna impedance, the real part and imaginary part can be calculated re-

spectively as

$$z_A = R_{z_A} + j \cdot X_{z_A}, \tag{B.1}$$

where $R_{z_A}$ is the resistance and $X_{z_A}$ is the reactance. The calculation of $R_{z_A}$ and $X_{z_A}$ is given by

$$\begin{aligned} R_{z_A} =& \frac{\eta}{2\pi \sin^2\left(\frac{KL}{2}\right)} \left\{ \gamma_0 + \ln(KL) - C_i(KL) + \frac{\sin(KL)}{2} \cdot [S_i(2KL) - 2S_i(KL)] \right. \\ & \left. + \frac{\cos(KL)}{2} \left[ \gamma_0 + \ln\left(\frac{KL}{2}\right) + C_i(2KL) - 2C_i(KL) \right] \right\}, \end{aligned} \tag{B.2.1}$$

$$\begin{aligned} X_{z_A} =& \frac{\eta}{4\pi \sin^2\left(\frac{KL}{2}\right)} \{ 2S_i(KL) + \cos(KL) \cdot [2S_i(KL) - S_i(KL)] \\ & - \sin(KL) \cdot \left[ 2C_i(KL) - C_i(2KL) - C_i\left(\frac{2Ka^2}{L}\right) \right] \}. \end{aligned} \tag{B.2.2}$$

where $\eta = 120\pi$ is the intrinsic impedance, $C_i$ and $S_i$ denote the cosine integral function and sine integral function, respectively. $\gamma_0$ is the Euler constant, $K = \frac{2\pi}{\lambda_0}$, and $L = l \cdot \lambda_0$, where $\lambda_0$ is the carrier wavelength and $l$ is the length of the dipole antennas normalised by the carrier wavelength. $a$ is the radius of the wire and a typical value $a = 0.001m$ is employed throughout this thesis. For typical half-wavelength dipole antennas, $l = 0.5$ and the resulting $z_A = (73 + j \cdot 42.5)\,\Omega$.

For a uniform linear half-wavelength dipole antenna array with side-by-side configuration, as employed in the simulations of this thesis, the mutual impedance can be calculated as

$$z_{M_k} = R_{m_k} + j \cdot X_{m_k}, \tag{B.3}$$

where the real part $R_{m_k}$ and imaginary part $X_{m_k}$ can be calculated respectively as

$$R_{m_k} = \frac{\eta}{4\pi} \{ 2C_i[u_0(k)] - C_i[u_1(k)] - C_i[u_2(k)] \}, \tag{B.4.1}$$

$$X_{m_k} = -\frac{\eta}{4\pi} \{ 2S_i[u_0(k)] - S_i[u_1(k)] - S_i[u_2(k)] \}, \tag{B.4.2}$$

where $u_0$, $u_1$ and $u_2$ can be calculated as

$$u_0(k) = 2\pi \cdot kd, \tag{B.5.1}$$

$$u_1(k) = 2\pi \cdot k\left(\sqrt{d^2+l^2}+l\right), \tag{B.5.2}$$

$$u_2(k) = 2\pi \cdot k\left(\sqrt{d^2+l^2}-l\right), \tag{B.5.3}$$

where $d$ is the normalised antenna spacing.

# Appendix C: BER over Correlated Rayleigh Channels

Under correlated Rayleigh channels, based on [30, 112, 113] the distribution of the SNR $\gamma_k$ on the $k$-th stream for ZF precoding is expressed as

$$f(\gamma_k) = \frac{\kappa_k^2 e^{-\frac{\gamma_k \kappa_k^2}{\gamma_0}}}{\gamma_0 (K-N_t)!}\left(\frac{\gamma_k \kappa_k^2}{\gamma_0}\right)^{K-N_t}, \tag{C.1}$$

where $\gamma_0$ is the transmit signal-to-noise ratio per bit ($E_b N_0$). When $K = N_t$, as assumed in the thesis, (C.1) can be simplified into

$$f_{K=N_t}(\gamma_k) = \frac{\kappa_k^2}{\gamma_0} e^{-\frac{\kappa_k^2}{\gamma_0}\cdot\gamma_k}. \tag{C.2}$$

Note that the conditional BER of the $k$-th stream can be expressed as [30]

$$P(\gamma_k) = Q\left(\sqrt{2\gamma_k}\right), \tag{C.3}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}dt$ is the $Q$-function. Then, the average BER for stream $k$ can be obtained by averaging $P(\gamma_k)$ over its probability density func-

tion, given by

$$\begin{aligned} P_k &= \int_0^\infty P(\gamma_k) f_{K=N_t}(\gamma_k) d\gamma_k \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi}} \int_{\sqrt{2\gamma_k}}^\infty e^{-\frac{t^2}{2}} dt \frac{\kappa_k^2}{\gamma_0} e^{-\frac{\kappa_k^2}{\gamma_0}} d\gamma_k \\ &= \frac{\kappa_k^2}{\sqrt{2\pi}\gamma_0} \int_0^\infty e^{-\frac{t^2}{2}} dt \int_0^{\frac{t^2}{2}} e^{-\frac{\kappa_k^2}{\gamma_0}} d\gamma_k \\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{\gamma_0+\kappa_k^2}{\gamma_o} \cdot \frac{t^2}{2}} dt \end{aligned} \tag{C.4}$$

As we have

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{t^2}{2}} dt = \frac{1}{2} \\ &\frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{\gamma_0+\kappa_k^2}{\gamma_o} \cdot \frac{t^2}{2}} dt = \frac{1}{2} \sqrt{\frac{\gamma_0}{\kappa_k^2 + \gamma_0}} \end{aligned} \tag{C.5}$$

the average BER for stream $k$ is obtained as

$$P_k = \frac{1}{2} \left[ 1 - \sqrt{\frac{\gamma_0}{\kappa_k^2 + \gamma_0}} \right]. \tag{C.6}$$

Finally, the average BER over all MIMO streams can be calculated as

$$P_b = \frac{1}{2K} \sum_{k=1}^{K} \left[ 1 - \sqrt{\frac{\gamma_0}{\kappa_k^2 + \gamma_0}} \right]. \tag{C.7}$$

# Appendix D: Coordinate Transformation

We employ the 8-PSK modulation in Fig. 8.2 as an example to demonstrate the coordinate transformation, where we focus on the constellation point 'A' in Fig. 8.2. Accordingly, in the conventional real-imaginary complex plane, for node 'B' in Fig. 8.2, we have

$$\vec{OB} = \mathbf{h}_k \mathbf{x}_T = B_r \cdot 1 + B_i \cdot j, \tag{D.1}$$

where 1 and $j$ are the bases, and we denote $(B_r, B_i)$ as the corresponding coordinate. Subsequently, $B_r$ and $B_i$ are obtained as

$$\begin{aligned} B_r &= \Re\left(\mathbf{h}_k \mathbf{x}_T\right) = \Re\left(\mathbf{h}_k\right) \mathbf{x}_T^{\Re} - \Im\left(\mathbf{h}_k\right) \mathbf{x}_T^{\Im}, \\ B_i &= \Im\left(\mathbf{h}_k \mathbf{x}_T\right) = \Im\left(\mathbf{h}_k\right) \mathbf{x}_T^{\Re} + \Re\left(\mathbf{h}_k\right) \mathbf{x}_T^{\Im}. \end{aligned} \tag{D.2}$$

In the plane expanded by the two detection thresholds that correspond to the constellation point 'A', $\vec{OB}$ is decomposed into

$$\vec{OB} = \mathbf{h}_k \mathbf{x}_T = \alpha_k^{\mathbb{A}} s_k^{\mathbb{A}} + \alpha_k^{\mathbb{B}} s_k^{\mathbb{B}}. \tag{D.3}$$

Based on the fact that $\alpha_k^{\mathbb{A}}$ and $\alpha_k^{\mathbb{B}}$ are real numbers, (D.3) is further transformed into

$$\begin{aligned} \mathbf{h}_k \mathbf{x}_T &= \alpha_k^{\mathbb{A}} \left( A_k^{\Re} + A_k^{\Im} \cdot j \right) + \alpha_k^{\mathbb{B}} \left( B_k^{\Re} + B_k^{\Im} \cdot j \right) \\ &= \left( A_k^{\Re} \alpha_k^{\mathbb{A}} + B_k^{\Re} \alpha_k^{\mathbb{B}} \right) + \left( A_k^{\Im} \alpha_k^{\mathbb{A}} + B_k^{\Im} \alpha_k^{\mathbb{B}} \right) \cdot j. \end{aligned} \tag{D.4}$$

By substituting (D.2) into (D.4), we obtain

$$\begin{aligned} B_r &= \Re\left(\mathbf{h}_k\right) \mathbf{x}_T^{\Re} - \Im\left(\mathbf{h}_k\right) \mathbf{x}_T^{\Im} = A_k^{\Re} \alpha_k^{\mathbb{A}} + B_k^{\Re} \alpha_k^{\mathbb{B}}, \\ B_i &= \Im\left(\mathbf{h}_k\right) \mathbf{x}_T^{\Re} + \Re\left(\mathbf{h}_k\right) \mathbf{x}_T^{\Im} = A_k^{\Im} \alpha_k^{\mathbb{A}} + B_k^{\Im} \alpha_k^{\mathbb{B}}, \end{aligned} \tag{D.5}$$

which leads to the expression of $\alpha_k^{\Re}$ and $\alpha_k^{\Im}$, given by

$$\begin{aligned} \alpha_k^{\mathbb{A}} &= \frac{B_k^{\Im} B_r - B_k^{\Re} B_i}{A_k^{\Re} B_k^{\Im} - A_k^{\Im} B_k^{\Re}} \\ &= \frac{B_k^{\Im} \left[ \mathbf{h}_k^{\Re} \mathbf{x}_T^{\Re} - \mathbf{h}_k^{\Im} \mathbf{x}_T^{\Im} \right] - B_k^{\Re} \left[ \mathbf{h}_k^{\Im} \mathbf{x}_T^{\Re} + \mathbf{h}_k^{\Re} \mathbf{x}_T^{\Im} \right]}{A_k^{\Re} B_k^{\Im} - A_k^{\Im} B_k^{\Re}} \\ &= \frac{B_k^{\Im} \mathbf{h}_k^{\Re} - B_k^{\Re} \mathbf{h}_k^{\Im}}{A_k^{\Re} B_k^{\Im} - A_k^{\Im} B_k^{\Re}} \mathbf{x}_T^{\Re} - \frac{B_k^{\Im} \mathbf{h}_k^{\Im} + B_k^{\Re} \mathbf{h}_k^{\Re}}{A_k^{\Re} B_k^{\Im} - A_k^{\Im} B_k^{\Re}} \mathbf{x}_T^{\Im}, \end{aligned} \tag{D.6}$$

and

$$\begin{aligned}\alpha_k^{\mathbb{B}} &= \frac{A_k^{\Re}B_i - A_k^{\Im}B_r}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}} \\ &= \frac{A_k^{\Re}\left[\mathbf{h}_k^{\Im}\mathbf{x}_T^{\Re} + \mathbf{h}_k^{\Re}\mathbf{x}_T^{\Im}\right] - A_k^{\Im}\left[\mathbf{h}_k^{\Re}\mathbf{x}_T^{\Re} - \mathbf{h}_k^{\Im}\mathbf{x}_T^{\Im}\right]}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}} \\ &= \frac{A_k^{\Re}\mathbf{h}_k^{\Im} - A_k^{\Im}\mathbf{h}_k^{\Re}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}}\mathbf{x}_T^{\Re} + \frac{A_k^{\Re}\mathbf{h}_k^{\Re} + A_k^{\Im}\mathbf{h}_k^{\Im}}{A_k^{\Re}B_k^{\Im} - A_k^{\Im}B_k^{\Re}}\mathbf{x}_T^{\Im}.\end{aligned} \tag{D.7}$$

The extension to other constellation points of 8-PSK and other PSK modulations can be similarly obtained and is omitted for brevity.

# Bibliography

[1] S. Sesia, M. Baker, and I. Toufik, *LTE - The UMTS Long Term Evolution*, 2nd ed. John Wiley and Sons, Aug. 2011.

[2] "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update," *White Paper*, pp. 2012-2017, 2013.

[3] D. Tse, and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.

[4] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems," *IEEE J. Sel. Areas. Commun.*, vol. 21, no. 3, pp. 281-302, Apr. 2003.

[5] L. Zheng, and D. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073-1096, May 2003.

[6] J. W. Wallace, and M. A. Jensen, "Mutual Coupling in MIMO Wireless Systems: A Rigorous Network Theory Analysis," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1317-1325, July 2004.

[7] Z. Li, Z. Du, M. Takahashi, K. Saito, and K. Ito, "Reducing Mutual Coupling of MIMO Antennas with Parasitic Elements for Mobile Terminals," *IEEE Trans. Ant. Propag.*, vol. 60, no. 2, pp. 473-481, Feb. 2012.

[8] A. Kalis, A. G. Kanatas, and C. B. Papadias, *Parasitic Antenna Arrays for Wireless MIMO Systems*. New York, NY, USA: Springer, 2014.

[9] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Trans. Wireless Commun.*,

vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[10] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Sig. Process. Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013.

[11] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid Beamforming for Massive MIMO: A Survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134-141, Sept. 2017.

[12] S. Henault, and Y. Antar, "Unifying the Theory of Mutual Coupling Compensation in Antenna Arrays," *IEEE Ant. Propag. Mag.*, vol. 57, no. 2, pp. 104-122, Apr. 2015.

[13] S. Ebadi, and A. Semnani, "Mutual Coupling Reduction in Waveguide-Slot-Array Antennas Using Electromagnetic Bandgap (EBG) Structures," *IEEE Ant. Propag. Mag.*, vol. 56, no. 3, pp. 68-79, June 2014.

[14] K. Gyoda, and T. Ohira, "Design of Electronically Steerable Passive Array Radiator (ESPAR) Antennas," in *IEEE Antennas and Propagation Society International Symposium. Transmitting Waves of Progress to the Next Millennium. 2000 Digest. Held in conjunction with: USNC/URSI National Radio Science Meeting*, vol. 2, Salt Lake City, UT, USA, 2000, pp. 922-925.

[15] A. Kalis, A. G. Kanatas, and C. B. Papadias, "A Novel Approach to MIMO Transmission Using a Single RF Front End," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 6, pp. 972-980, Aug. 2008.

[16] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of One-Bit Quantized Precoding for the Multiuser Massive MIMO Downlink," *IEEE Trans. Sig. Process.*, vol. 65, no. 17, pp. 4624-4634, Sept. 2017.

[17] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication-Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, pp. 537-544, Mar. 2005.

[18] R. Chen, C. Li, J. Li, and Y. Zhang, "Low Complexity User Grouping Vector Perturbation," *IEEE Wireless Commun. Lett.*, vol. 1, no. 3, pp. 189-192, June 2012.

[19] I. Salonen, A. Toropainen, and P. Vainikainen, "Linear Pattern Correction in a Small Microstrip Antenna Array," *IEEE Trans. Ant. Propag.*, vol. 52, no. 2, pp. 578-586, Feb. 2004.

[20] G. Moreno, H. M. Bernety, and A. B. Yakovkev, "Reduction of Mutual Coupling between Strip Dipole Antennas at Terahertz Frequencies with an Elliptically Shaped Graphene Monolayer," *IEEE Ant. Wireless Propag. Lett.*, vol. 15, pp. 1533-1536, Dec. 2015.

[21] V. I. Barousis, C. B. Papadias, and R. R. Müller, "A New Signal Model for MIMO Communication with Compact Parasitic Arrays," in *2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, Athens, 2014, pp. 109-113.

[22] G. C. Alexandropoulos, V. I. Barousis, and C. B. Papadias, "Precoding for Multiuser MIMO Systems with Single-Fed Parasitic Antenna Arrays," in *2014 IEEE Global Communications Conference*, Austin, TX, USA, 2014, pp. 3897-3902.

[23] S. Han, C. I. I, and C. Rowell, "Large-Scale Antenna Systems with Hybrid Analog and Digital Beamforming for Millimeter Wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186-194, Jan. 2015.

[24] A. Alkhateeb, O. EI Ayach, G. Leus, and R. W. Heath, "Channel Estimation and Hybrid Precoding for Millimeter Wave Cellular Systems." *IEEE J. Sel. Topics Sig. Process.*, vol. 8, no. 5, pp. 831-846, Oct. 2014.

[25] C. Mollen, J. Choi, E. G. Larsson, and R. W. Heath, "Uplink Performance of Wideband Massive MIMO with One-Bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 87-100, Oct. 2016.

[26] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput Analysis of Massive MIMO Uplink with Low-Resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038-4051,

June 2017.

[27] A. Mezghani, R. Ghiat, and J. A. Nossek, "Transmit Processing with Low Resolution D/A Converters," in *2009 16th IEEE International Conference on Electronics, Circuits and Systems - (ICECS 2009)*, Yasmine Hammamet, 2009, pp. 683-686.

[28] O. Castaneda, T. Goldstein, and C. Studer, "POKEMON: A Non-Linear Beamforming Algorithm for 1-Bit Massive MIMO," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, USA, 2017, pp. 3464-3468.

[29] I. S. Gradshteyn, and I. M. Ryzhik, *Table of Integral, Series and Products*, 6th ed. Academic Press, 2000.

[30] R. Xu, and F. C. M. Lau, "Performance Analysis for MIMO Systems Using Zero Forcing Detector over Fading Channels," *IEE Proceedings - Commun.*, vol. 153, no. 1, pp. 74-80, Feb. 2006.

[31] Y. Jiang, M. K. Varanasi, and J. Li, "Performance Analysis of ZF and MMSE Equalizers for MIMO Systems: An In-Depth Study of High SNR Regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2008-2026, Apr. 2011.

[32] B. Vucetic, and J. Yuan, *Space-Time Coding*, 1st ed. John Wiley and Sons, 2003.

[33] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A Vector-Perturbation Technique for Near-Capacity Multiantenna Multiuser Communication-Part I: Channel Inversion and Regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195-202, Jan. 2005.

[34] J. Maurer, J. Jalden, D. Seethaler, and G. Matz, "Vector Perturbation Precoding Revisited," *IEEE Trans. Sig. Process.*, vol. 59, no. 1, pp. 315-328, Jan. 2011.

[35] C. Masouros, M. Sellathurai, and T. Ratnarajah, "A Low-Complexity Sequential Encoder for Threshold Vector Perturbation," *IEEE Commun. Lett.*, vol. 17, no. 12, pp. 2225-2228, Dec. 2013.

[36] ——, "Computationally Efficient Vector Perturbation Precoding Using Thresholded Optimizations," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 1880-1890, May 2013.

[37] ——, "Maximizing Energy Efficiency in the Vector Precoded MU-MISO Downlink by Selective Perturbation," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4974-4984, Sept. 2014.

[38] ——, "Vector Perturbation based on Symbol Scaling for Limited Feedback MISO Downlinks," *IEEE Trans. Sig. Process.*, vol. 62, no. 3, pp. 562-571, Feb. 2014.

[39] M. Bengtsson, and B. Ottersten, "Optimal and Sub-Optimal Transmit Beamforming," *Handbook of Antennas in Wireless Communications*, Jan. 2001.

[40] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Linear Precoding via Conic Optimization for Fixed MIMO Receivers," *IEEE Trans. Sig. Process.*, vol. 54, no. 1, pp. 161-176, Jan. 2006.

[41] C. Masouros, and G. Zheng, "Exploiting Known Interference as Green Signal Power for Downlink Beamforming Optimization," *IEEE Trans. Sig. Process.*, vol. 63, no. 14, pp. 3628-3640, July 2015.

[42] M. Joham, W. Utschick, and J. A. Nossek, "Linear Transmit Processing in MIMO Communication Systems," *IEEE Trans. Sig. Process.*, vol. 53, no. 8, pp. 2700-2712, Aug. 2005.

[43] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels," *IEEE Trans. Sig. Process.*, vol. 52, no. 2, pp. 461-471, Feb. 2004.

[44] T. Haustein, C. von Helmolt, E. Jorswieck, V. Jungnickel, and V. Pohl, "Performance of MIMO Systems with Channel Inversion," in *IEEE 55th Vehicular Technology Conference, VTC Spring 2002 (Cat. No.02CH37367)*, vol. 1, 2002, pp. 35-39.

[45] M. Costa, "Writing on Dirty Paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439-441, May 1983.

[46] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Interference Optimization for Transmit Power Reduction in Tomlinson-Harashima Precoded MIMO Downlinks," *IEEE Trans. Sig. Process.*, vol. 60, no. 5, pp. 2470-2481, May 2012.

[47] C. Windpassinger, R. F. H. Fischer, T. Vencel, and J. B. Huber, "Precoding in Multiantenna and Multiuser Communications," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1305-1316, July 2004.

[48] C. B. Chae, S. Shim, and R. W. Heath, "Block Diagonalized Vector Perturbation for Multiuser MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4051-4057, Nov. 2008.

[49] M. Schubert, and H. Boche, "Solution of the Multiuser Downlink Beamforming Problem with Individual SINR Constraints," *IEEE Trans. Veh. Tech.*, vol. 53, no. 1, pp. 18-28, Jan. 2004.

[50] C. Masouros, T. Ratnarajah, M. Sellathurai, C. B. Papadias, and A. K. Shukla, "Known Interference in the Cellular Downlink: A Performance Limiting Factor or a Source of Green Signal Power?" *IEEE Commun. Mag.*, vol. 51, no. 10, pp. 162-171, Oct. 2013.

[51] G. Zheng, I. Krikidis, C. Masouros, S. Timotheou, D. A. Toumpakaris, and Z. Ding, "Rethinking the Role of Interference in Wireless Networks," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 152-158, Nov. 2014.

[52] C. Masouros, and E. Alsusa, "Two-Stage Transmitter Precoding based on Data-Driven Code Hopping and Partial Zero Forcing Beamforming for MC-DCMA Communications," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3634-3645, July 2009.

[53] ——, "Soft Linear Precoding for the Downlink of DS/CDMA Communication Systems," *IEEE Trans. Veh. Tech.*, vol. 59, no. 1, pp. 203-215, Jan. 2010.

[54] ——, "Dynamic Linear Precoding for the Exploitation of Known Interference in MIMO Broadcast Systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1396-1404, Mar. 2009.

[55] C. Masouros, "Correlation Rotation Linear Precoding for MIMO Broadcast Systems," *IEEE Trans. Sig. Process.*, vol. 59, no. 1, pp. 252-262, Jan. 2011.

[56] R. B. Ertel, P. Cardieri, K. W. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of Spatial Channel Models for Antenna Array Communication Systems," *IEEE Personal Commun.*, vol. 5, no. 1, pp. 10-22, 1998.

[57] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading Correlation and Its Effect on the Capacity of Multielement Antenna Systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502-513, Mar. 2009.

[58] A. M. Sayeed, "Deconstructing Multiantenna Fading Channels," *IEEE Trans. Sig. Process.*, vol. 50, no. 10, pp. 2563-2579, Oct. 2002.

[59] M. Steinbauer, A. Molisch, and E. Bonek, "The Double-Directional Radio Channel," *IEEE Ant. Propag. Mag.*, vol. 43, no. 4, pp. 51-63, Aug. 2001.

[60] A. Forenza, D. J. Love, and R. W. Heath, "Simplified Spatial Correlation Models for Clustered MIMO Channels with Different Array Configurations," *IEEE Trans. Veh. Tech.*, vol. 56, no. 4, pp. 1924-1934, July 2007.

[61] C. Masouros, J. Chen, K. Tong, M. Sellathurai, T. Ratnarajah, and J. Wang, "Large Scale Antenna Arrays with Increasing Antennas in Limited Physical Space," *Commun., China*, vol. 11, pp. 7-15, Nov. 2014.

[62] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Large-Scale MIMO Transmitters in Fixed Physical Spaces: The Effect of Transmit Correlation and Mutual Coupling," *IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2794-2804, July 2013.

[63] C. A. Balanis, *Antenna Theory: Analysis and Design*, 4th ed. Wiley-Blackwell, 2016.

[64] B. Clerckx, C. Craeye, D. V-Janvier, and C. Oestges, "Impact on Antenna Coupling on 2×2 MIMO Communications," *IEEE Trans. Veh. Tech.*, vol. 56, no. 3, pp. 1009-1018, May 2007.

[65] Z. Pi, and F. Khan, "An Introduction to Millimeter-Wave Mobile Broadband Systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101-107, June 2011.

[66] R. W. Heath, M. A. Gonzalez, S. Rangan, W. Roh, and A. M. Sayeed, "An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems," *IEEE J. Sel. Topics Sig. Process.*, vol. 10, no. 3, pp. 436-453, Apr. 2016.

[67] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited Feedback Hybrid Precoding for Multi-User Millimeter Wave Systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481-6494, July 2015.

[68] M. B. Shenouda, and T. N. Davidson, "On the Design of Linear Transceivers for Multiuser Systems with Channel Uncertainty," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 6, pp. 1015-1024, Aug. 2008.

[69] Q. Zhang, C. He, and L. Jiang, "Per-Stream MSE based Linear Transceiver Design for MIMO Interference Channels with CSI Error," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1676-1689, May 2015.

[70] M. Ding, and S. Blostein, "MIMO Minimum Total MSE Transceiver Design with Imperfect CSI at Both Ends," *IEEE Trans. Sig. Process.*, vol. 57, no. 3, pp. 1141-1150, Mar. 2009.

[71] S. Lu, H. T. Hui, M. E. Bialkowski, X. Liu, H. S. Hui, and N. V. Shuley, "Effects on Antenna Mutual Coupling on the Performance of MIMO Systems," in *2009 Symposium on Information Theory*, Benelux, May 2008, pp. 2945-2948.

[72] I. Gupta, and A. Ksienski, "Effect of Mutual Coupling on the Performance of Adaptive Arrays," *IEEE Trans. Ant. Propag.*, vol. 31, no. 5, pp. 785-791, Sept. 1983.

[73] A. A. Abouda, and S. G. Haggman, "Effect of Mutual Coupling on Capacity of MIMO Wireless Channels in High SNR," *Progress in Electromagnetics Research*, vol. 65, pp. 27-40, 2005.

[74] H. Steyskal, and J. S. Herd, "Mutual Coupling Compensation in Small Array Antennas," *IEEE Trans. Ant. Propag.*, vol. 38, no. 2, pp. 1971-1975, Dec. 1990.

[75] J. Corcoles, M. A. Gonzalez, and J. Rubio, "Mutual Coupling Compensation in Arrays Using a Spherical Wave Expansion of the Radiated Field," *IEEE Ant. Wireless Propag. Lett.*, vol. 8, pp. 108-111, Jan. 2009.

[76] J. Rubio, J. F. Izquierdo, and J. Corcoles, "Mutual Coupling Compensation Matrices for Transmitting and Receiving Arrays," *IEEE Trans. Ant. Propag.*, vol. 63, no. 2, pp. 839-843, Feb. 2015.

[77] S. Farsi, H. Aliakbarian, D. Schreurs, B. Nauwelaers, and G. A. E. Vandenbosch, "Mutual Coupling Reduction between Planar Antennas by Using a Simple Microstrip U-Section," *IEEE Ant. Wireless Propag. Lett.*, vol. 11, pp. 1501-1503, 2012.

[78] V. I. Barousis, and C. B. Papadias, "Arbitrary Precoding with Single-Fed Parasitic Arrays: Closed-Form Expression and Design Guidelines," *IEEE Wireless Commun. Lett.*, vol. 3, no. 2, pp. 229-232, Apr. 2014.

[79] B. Han, V. I. Barousis, C. B. Papadias, A. Kalis, and R. Prasad, "MIMO over ESPAR with 16-QAM Modulation," *IEEE Wireless Commun. Lett.*, vol. 2, no. 6, pp. 687-690, Apr. 2013.

[80] O. EI Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially Sparse Precoding in Millimeter Wave MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499-1513, Mar. 2014.

[81] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, A. Alkhateeb, and R. W. Heath, "Hybrid MIMO Architectures for Millimeter Wave Communications: Phase Shifters or Switches?" *IEEE Access*, vol. 4, pp. 247-267, Jan. 2016.

[82] Y. Zeng, and R. Zhang, "Millimeter Wave MIMO with Lens Antenna Array: A New Path Division Multiplexing Paradigm," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1557-1571, Apr. 2016.

[83] Y. Gao, M. Khaliel, F. Zheng, and T. Kaiser, "Rotman Lens based Hybrid Analog-Digital Beamforming in Massive MIMO Systems: Array Architectures, Beam Selection Algorithms and Experiments," *IEEE Trans. Veh. Tech.*, vol. 66, no. 10, pp. 9134-9148, Oct. 2017.

[84] A. Li, and C. Masouros, "A Constellation Scaling Approach to Vector Perturbation for Adaptive Modulation in MU-MIMO," *IEEE Wireless Commun. Lett.*, vol. 4, no. 3, pp. 289-292, June 2015.

[85] ——, "A Two-Stage Vector Perturbation Scheme for Adaptive Modulation in Downlink MU-MIMO," *IEEE Trans. Veh. Tech.*, vol. 65, no. 9, pp. 7785-7791, Sept. 2016.

[86] H. S. Han, S. H. Park, S. Lee, and I. Lee, "A Decoupling Approach for Low-Complexity Vector Perturbation in Multiuser Downlink Systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1697-1701, June 2011.

[87] G. H. Golub, and C. F. V. Load, *Matrix Computations*, 3rd ed. The John Hopkins University Press, 1989.

[88] H. Sung, S. C. Lee, and I. Lee, "Generalized Channel Inversion Methods for Multiuser MIMO Systems," *IEEE Trans. Commun.*, vol. 57, no. 11, pp. 3489-3499, Nov. 2009.

[89] A. Li, and C. Masouros, "Mutual Coupling Exploitation for Point-to-point MIMO by Constructive Interference," in *2017 IEEE International Conference on Communications (ICC)*, Paris, France, 2017, pp. 1-6.

[90] ——, "Exploiting Constructive Mutual Coupling in P2P MIMO by Analog-Digital Phase Alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1948-1962, Mar. 2017.

[91] S. Biswas, C. Masouros, and T. Ratnarajah, "Performance Analysis of Large Multi-User MIMO Systems with Space-Constrained 2-D Antenna Arrays," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3492-3505, May 2016.

[92] D. Piazza, N. J. Kirsch, A. Forenza, R. W. Heath, and K. R. Dandekar, "Design and Evaluation of a Reconfigurable Antenna Array for MIMO Systems," *IEEE Trans. Ant. Propag.*, vl. 56, no. 3, pp. 869-881, Mar. 2008.

[93] P. L. Kafle, A. Intarapanich, A. B. Sesay, J. McRory, and R. J. Davies, "Spatial Correlation and Capacity Measurements for Wideband MIMO Channels in Indoor Office Environment," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1560-1571, May 2008.

[94] G. Alfano, A. M. Tulino, A. Lozano, and S. Verdu, "Capacity of MIMO Channels with One-Sided Correlation," in *Proc. 2004 IEEE Conf. Spread Spectrum Techniques and Applications*, 2004, pp. 515-519.

[95] A. M. Tulino, A. Lozano, and S. Verdu, "Impact of Antenna Correlation on the Capacity of Multiantenna Channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491-2509, July 2005.

[96] H. Liu, Y. Song, and R. C. Qiu, "The Impact of Fading Correlation on the Error Performance of MIMO Systems over Rayleigh Fading Environments," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2014-2019, Sept. 2005.

[97] C.-K. Wen, Y.-N. Lee, J.-T. Chen, and P. Ting, "Asymptotic Spectral Efficiency of MIMO Multiple-Access Wireless Systems Exploring Only Channel Spatial Correlations," *IEEE Trans. Sig. Process.*, vol. 53, no. 6, pp. 2059-2073, June 2005.

[98] A. Alexiou, and M. Qaddi, "Robust Linear Precoding to Compensate for Antenna Correlation in Orthogonal Space-Time Block Coded Systems," in *Sensor Array and Multichannel Signal Processing Workshop Proceedings*, pp. 701-705, 2004.

[99] H. R. Bahrami, and T. Le-Ngoc, "Precoder Design based on the Channel Correlation Matrices," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 2579-2587, Dec. 2006.

[100] J. Akhtar, and D. Gesbert, "Spatial Multiplexing over Correlated

MIMO Channels with a Closed-Form Precoder," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2400-2409, Sept. 2005.

[101] J. W. Wallace, and M. A. Jensen, "The Capacity of MIMO Wireless Systems with Mutual Coupling," in *Proceedings IEEE 56th Vehicular Technology Conference*, vol. 2, 2002, pp. 696-700.

[102] V. Jungnickel, V. Pohl, and C. von Helmolt, "Capacity of MIMO Systems with Closely Spaced Antennas," *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 361-363, Aug. 2003.

[103] X. Liu, and M. E. Bialkovwski, "Effect of Antenna Mutual Coupling on MIMO Channel Estimation and Capacity," *International Journal of Antennas and Propagation*, Article ID 306173, 2010.

[104] T. Svantesson, and A. Ranheim, "Mutual Coupling Effects on the Capacity of Multielement Antenna Systems," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, 2001, pp. 2485-2488.

[105] H. T. Hui, "A Practical Approach to Compensate for the Mutual Coupling Effect in an Adaptive Dipole Array," *IEEE Trans. Ant. Propag.*, vol. 52, no. 5, pp. 1262-1269, May 2004.

[106] R. S. Adve, and T. K. Sarkar, "Compensation for the Effects of Mutual Coupling on Direct Data Domain Adaptive Algorithms," *IEEE Trans. Ant. Propag.*, vol. 48, no. 1, pp. 86-94, Jan. 2000.

[107] J. Choma, and W. K. Chen, *Feedback Networks: Theory and Circuit Applications*. NJ, USA: World Scientific Publishing, May 2007.

[108] H. A. Haus, *Electromagnetic Noise and Quantum Optical Measurements*. New York, NY, USA: Springer, Nov. 2000.

[109] M. D. Migliore, "An Intuitive Electromagnetic Approach to MIMO Communication Systems," *IEEE Ant. Propag. Mag.*, vol. 48, no. 3, pp. 128-137, June 2006.

[110] P. N. Vasileiou, K. Maliatsos, E. D. Thomatos, and A. G. Kanatas, "Reconfigurable Orthonormal Basis Patterns Using ESPAR Antennas,"

*IEEE Ant. Wireless Propag. Lett.*, vol. 12, pp. 448-451, 2013.

[111] Y. Zhou, R. S. Adve, and S. V. Hum, "Design and Evaluation of Pattern Reconfigurable Antennas for MIMO Applications," *IEEE Trans. Ant. Propag.*, vol. 62, no. 3, pp. 1084-1092, Oct. 2013.

[112] D. Gore, R. W. Heath, and A. Paulraj, "On Performance of the Zero Forcing Receiver in Presence of Transmit Correlations," in *Proceedings IEEE International Symposium on Information Theory*, 2002, pp. 195.

[113] M. Kiessling, and J. Speidel, "Analytical Performance of MIMO Zero-Forcing Receivers in Correlated Rayleigh Fading Environments," in *2003 4th IEEE Workshop on Signal Processing Advances in Wireless Communications - SPAWC 2003 (IEEE Cat. No.03EX689)*, 2003, pp. 383-387.

[114] J.-S. Fu, "Adaptive Impedance Matching Circuits based on Ferroelectric and Semiconductor Varactors," Ph.D. dissertation, University of Michigan, 2009.

[115] R. York, A. Nagra, E. Erker, T. Taylor, P. Periaswamy, J. Speck, S. Streiffer, and O. Auciello, "Microwave Intergrated Circuits Using Thin-Film BST," in *Proceedings of the 2000 12th IEEE International Symposium on Applications of Ferroelectrics (ISAF 2000)*, vol. 1, Honolulu, HI, 2000, pp. 195-200.

[116] J.-S. Fu, X. A. Zhu, J. D. Phillips, and A. Mortazawi, "Improving Linearity of Ferroelectric-based Microwave Tunable Circuits," *IEEE Trans. Microwave Theory and Techniques*, vol. 55, no. 2, pp. 354-360, Feb. 2007.

[117] O. N. Alrabadi, C. Divarathne, P. Tragas, A. Kalis, N. Marchetti, C. B. Papadias, and R. Prasad, "Spatial Multiplexing with a Single Radio: A Proof-of-Concept Experiments in an Indoor Environment with a 2.6GHz Prototype," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 178-180, Feb. 2011.

[118] A. Li, and C. Masouros, "Exploiting Mutual Coupling by Means of Analog-Digital Zero Forcing," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA,

2017, pp. 3564-3568.

[119] ——, "Analog-Digital Beamforming for MIMO with Tunable Loads by Mutual Coupling Exploitation," in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Sapporo, Japan, 2017, pp. 1-5.

[120] A. Li, C. Masouros, and M. Sellathurai, "Analog-Digital Beamforming in the MU-MISO Downlink by Use of Tunable Antenna Loads," *IEEE Trans. Veh. Tech.*, DOI:10.1109/TVT.2017.2776563, 2017.

[121] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A Unified Convergence Analysis of Block Successive Minimization Methods for Nonsmooth Optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126-1153, 2013.

[122] F. Wang, X. Wang, and Y. Zhu, "Transmit Beamforming for Multiuser Downlink with Per-Antenna Power Constraints," in *2014 IEEE International Conference on Communications (ICC)*, Sydney, NSW, 2014, pp. 4692-4697.

[123] M. F. Hanif, L.-N. Tran, A. Tolli, and M. Juntti, "Computationally Efficient Robust Beamforming for SINR Balancing in Multicell Downlink with Applications to Large Antenna Array Systems," *IEEE Trans. Commun.*, vol. 62, no. 6, pp. 1908-1920, June 2014.

[124] S. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

[125] E. Song, Q. Shi, M. Sanjabi, R. Sun, and Z.-Q. Luo, "Robust SINR-Constrained MISO Downlink Beamforming: When is Semidefinite Relaxation Tight?" in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 3096-3099.

[126] Z.-Q. Luo, W. Ma, A. So, Y. Ye, and S. Zhang, "Semidefinite Relaxation of Quadratic Optimization Problems," *IEEE Sig. Process. Mag.*, vol. 27, no. 3, pp. 20-34, May 2010.

[127] Y. Huang, and D. Palomar, "Rank-Constrained Separable Semidefinite Programming with Applications to Optimal Beamforming," *IEEE*

*Trans. Sig. Process.*, vol. 58, no. 2, pp. 664-678, Feb. 2010.

[128] A. Li, and C. Masouros, "Performance Analysis for Single-Fed ESPAR in the Presence of Impedance Errors and Imperfect CSI," in *2016 IEEE International Conference on Communications (ICC)*, Kuala Lumpur, 2016, pp. 1-6.

[129] A. Li, C. Masouros, and C. B. Papadias, "MIMO Transmission for Single-Fed ESPAR with Quantized Loads," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 2863-2876, July 2017.

[130] A. Li, C. Masouros, M. Sellathurai, and C. B. Papadias, "Tunable Load MIMO with Quantized Loads," in *25th European Signal Processing Conference (EUSIPCO)*, Kos, Greece, 2017, pp. 1-6.

[131] B. Han, V. I. Barousis, A. Kalis, C. B. Papadias, A. G. Kanatas, and R. Prasad, "A Single RF MIMO Loading Network for High Order Modulation Schemes," *International Journal of Antennas and Propagation*, vol. 3, pp. 1-10, July 2014.

[132] L. Zhou, F. A. Khan, T. Ratnarajah, and C. B. Papadias, "Achieving Arbitrary Signals Transmission Using a Single Radio Frequency Chain," *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4865-4878, Dec. 2015.

[133] M. A. Sedaghat, V. I. Barousis, R. R. Muller, and C. B. Papadias, "Load Modulated Arrays: A Low-Complexity Antenna," *IEEE Commun. Mag.*, vol. 54, no. 3, pp. 46-52, Mar. 2016.

[134] D. J. R. Chisaguano, Y. Hou, T. Higashino, and M. Okada, "Low-Complexity Channel Estimation and Detection for MIMO-OFDM Receiver with ESPAR Antenna," *IEEE Trans. Veh. Tech.*, vol. 65, no. 10, pp. 8297-8308, Oct. 2016.

[135] D. S. Holder, Ed., *Electrical Impedance Tomography: Methods, History and Applications.* FL, USA: CRC Press, Dec. 2004.

[136] S. L. Carson, M. E. Orazem, O. D. Crisalle, and L. Garcia-Rubio, "On the Error Structure of Impedance Measurements - Simulation of FRA

Instrumentation," *Journal of the Electrochemical Society*, vol. 150, no. 10, pp. 477-490, 2003.

[137] ——, "On the Error Structure of Impedance Measurements - Series Expansions," *Journal of the Electrochemical Society*, vol. 150, no. 10, pp. 501-511, 2003.

[138] G. L. Studer, *Principles of Mobile Communications*, 3rd ed. New York, NY, USA: Springer, 2011.

[139] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. Zhang, "What Will 5G Be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065-1082, June 2014.

[140] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335-349, May 2013.

[141] V. Venkateswaran, and A. van de Veen, "Analog Beamforming in MIMO Communications with Phase Shift Networks and Online Channel Estimation," *IEEE Trans. Sig. Process.*, vol. 58, no. 8, pp. 4131-4143, Apr. 2010.

[142] F. Gholam, J. Via, and I. Santamaria, "Beamforming Design for Simplified Analog Antenna Combining Architectures," *IEEE Trans. Veh. Tech.*, vol. 60, no. 5, pp. 2373-2378, Apr. 2011.

[143] A. Garcia, V. Venkateswaran, P. Rulikowski, and C. Masouros, "Hybrid Analog-Digital Precoding Revisited under Realistic RF Modeling," *IEEE Wireless Commun. Lett.*, vol. 5, no. 5, pp. 528-531, Oct. 2016.

[144] L. Liang, W. Xu, and X. Dong, "Low-Complexity Hybrid Precoding in Massive Multiuser MIMO Systems," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 653-656, Dec. 2014.

[145] D. W. K. Ng, E. S. Lo, and R. Schober, "Wireless Information and Power Transfer: Energy Efficiency Optimization in OFDMA Systems," *IEEE*

*Trans. Wireless Commun.*, vol. 12, no. 2, pp. 6352-6370, Dec. 2013.

[146] T. Chen, Y. Yang, H. Zhang, H. Kim, and K. Horneman, "Network Energy Saving Technologies for Green Wireless Access Networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 30-38, Oct. 2011.

[147] K. Pentikousis, "In Search of Energy-Efficient Mobile Networking," *IEEE Commun. Mag.*, vol. 48, no. 1, pp. 95-103, Jan. 2010.

[148] S. Sudevalayam, and P. Kulkarni, "Energy Harvesting Sensor Nodes: Survey and Implications," *IEEE Commun. Surveys Tut.*, vol. 13, no. 3, pp. 443-461, 3rd Quart. 2011.

[149] G. Monti, L. Corchia, and L. Tarricone, "UHF Wearable Rectenna on Textile Materials," *IEEE Trans. Ant. Propag.*, vol. 61, no. 7, pp. 3869-3873, July 2013.

[150] U. Olgun, C.-C. Chen, and J. L. Volakis, "Investigation of Rectenna Array Configurations for Enhanced RF Power Harvesting," *IEEE Ant. Wireless Propag. Lett.*, vol. 10, pp. 262-265, 2011.

[151] W. Ejaz, M. Naeem, M. Basharat, A. Anpalagan, and S. Kandeepan, "Efficient Wireless Power Transfer in Software-Defined Wireless Sensor Networks," *IEEE Sensors Journal*, vol. 16, no. 20, pp. 7409-7420, Oct. 2016.

[152] T. He, K.-W. Chin, and S. Soh, "On Wireless Power Transfer and Max Flow in Rechargeable Wireless Sensor Networks," *IEEE Access*, vol. 4, pp. 4155-4167, Aug. 2016.

[153] Z. Wang, L. Duan, and R. Zhang, "Adaptively Directional Wireless Power Transfer for Large-Scale Sensor Networks," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1785-1800, May 2016.

[154] Y.-W. P. Hong, T.-C. Hsu, and P. Chennakesavula, "Wireless Power Transfer for Distributed Estimation in Wireless Passive Sensor Networks," *IEEE Trans. Sig. Process.*, vol. 64, no. 20, pp. 5382-5395, Oct. 2016.

[155] S. Bi, Y. Zeng, and R. Zhang, "Wireless Powered Communication Networks: An Overview," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 10-18, May 2016.

[156] V. Chawla, and D. S. Ha, "An Overview of Passive RFID," *IEEE Commun. Mag.*, vol. 45, no. 9, pp. 11-17, Sept. 2007.

[157] L. Varshney, "Transporting Information and Energy Simultaneously," in *2008 IEEE International Symposium on Information Theory*, Toronto, ON, 2008, pp. 1612-1616.

[158] P. Grover, and A. Sahai, "Shannon Meets Tesla: Wireless Information and Power Transfer," in *2010 IEEE International Symposium on Information Theory*, Austin, TX, 2010, pp. 2363-2367.

[159] R. Zhang, and C. K. Ho, "MIMO Broadcasting for Simultaneous Wireless Information and Power Transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989-2001, May 2013.

[160] Z. Ding, C. Zhong, D. W. K. Ng, M. Peng, H. A. Suraweera, R. Schober, and H. V. Poor, "Application of Smart Antenna Technologies in Simultaneous Wireless Information and Power Transfer," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 86-93, Apr. 2015.

[161] T. L. Thanh, M. D. Renzo, and J. P. Coon, "MIMO Cellular Networks with Simultaneous Wireless Information and Power Transfer," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Signal Processing (SPAWC)*, Edinburgh, 2016, pp. 1-5.

[162] M. Sheng, L. Wang, X. Wang, Y. Zhang, C. Xu, and J. Li, "Energy Efficient Beamforming in MISO Heterogeneous Cellular Networks with Wireless Information and Power Transfer," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 954-968, Apr. 2016.

[163] A. Ozcelikkale, T. McKelvey, and M. Viberg, "Wireless Information and Power Transfer in MIMO Channels under Rician Fading," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, QLD, 2015, pp. 3187-3191.

[164] H. Son, and B. Clerckx, "Joint Beamforming Design for Multi-User Wireless Information and Power Transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6397-6409, Aug. 2014.

[165] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO Beamforming for Simultaneous Wireless Information and Power Transfer," *IEEE Trans. Sig. Process.*, vol. 62, no. 18, pp. 4798-4810, Sept. 2014.

[166] S. Timotheou, G. Zheng, C. Masouros, and I. Krikidis, "Exploiting Constructive Interference for Simultaneous Wireless Information and Power Transfer in Multiuser Downlink Systems," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 5, pp. 1772-1784, May 2016.

[167] Z. Zong, H. Feng, F. R. Yu, N. Zhao, T. Yang, and B. Hu, "Optimal Transceiver Design for SWIPT for K-User MIMO Interference Channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 430-445, Jan. 2016.

[168] S. Lee, L. Liu, and R. Zhang, "Collaborative Wireless Energy and Information Transfer in Interference Channel," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 545-557, Sept. 2015.

[169] J. Park, and B. Clerckx, "Joint Wireless Information and Energy Transfer in a K-User MIMO Interference Channel," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5781-5796, Oct. 2014.

[170] J. Xiao, C. Xu, X. Huang, and J. Qin, "Robust Transceiver Design for Two-User MIMO Interference Channel with Simultaneous Wireless Information and Power Transfer," *IEEE Trans. Veh. Tech.*, vol. 65, no. 5, pp. 3823-3828, May 2016.

[171] X. Chen, D. W. K. Ng, and H.-H. Chen, "Secrecy Wireless Information and Power Transfer: Challenges and Opportunities," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 54-61, May 2016.

[172] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure Beamforming for MIMO Broadcasting with Wireless Information and Power Transfer," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2841-2853, Jan. 2015.

[173] X. Chen, J. Chen, and T. Liu, "Secure Transmission in Wireless Powered Massive MIMO Relaying Systems: Performance Analysis and Optimization," *IEEE Trans. Veh. Tech.*, vol. 65, no. 10, pp. 8025-8035, Oct. 2016.

[174] S. Wang, and B. Wang, "Robust Secure Transmit Design in MIMO Channels with Simultaneous Wireless Information and Power Transfer," *IEEE Sig. Process. Lett.*, vol. 22, no. 11, pp. 2147-2151, Nov. 2015.

[175] K. Yang, Q. Yu, S. Leng, B. Fan, and F. Wu, "Data and Energy Integrated Communication Networks for Wireless Big Data," *IEEE Access*, vol. 4, pp. 713-723, Feb. 2016.

[176] Y. Zeng, B. Clerckx, and R. Zhang, "Communications and Signals Design for Wireless Power Transmission," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2264-2290, May 2017.

[177] R. Zhang, and Y.-C. Liang, "Exploiting Multi-Antennas for Opportunistic Spectrum Sharing in Cognitive Radio Networks," *IEEE J. Sel. Areas Commun.*, vol. 2, no. 1, pp. 88-102, Feb. 2008.

[178] A. Ben-Tal, and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2001.

[179] A. Li, C. Masouros, F. Liu, and A. L. Swindlehurst, "Massive MIMO 1-Bit DAC Transmission: A Low-Complexity Symbol Scaling Approach," *arXiv preprint*, available online: http://arxiv.org/abs/1709.08278, 2017.

[180] A. Li, C. Masouros, and A. L. Swindlehurst, "1-Bit Massive MIMO Downlink Based on Constructive Interference," submitted to *2018 26th European Signal Processing Conference (EUSIPCO)*.

[181] A. Li, C. Masouros, and F. Liu, "Hybrid Analog-Digital 1-Bit DAC Transmission: From MMSE to Interference Exploitation," submitted to *2018 26th European Signal Processing Conference (EUSIPCO)*.

[182] R. H. Walden, "Analog-to-Digital Converter Survey and Analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539-550, Apr. 1999.

[183] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, and A. L. Swindlehurst, "Channel Estimation and Performance Analysis of One-Bit Massive MIMO Systems," *IEEE Trans. Sig. Process.*, vol. 65, no. 15, pp. 4075-4089, Aug. 2017.

[184] O. B. Usman, H. Jedda, A. Mezghani, and J. A. Nossek, "MMSE Precoder for Massive MIMO Using 1-Bit Quantization," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, 2016, pp. 3381-3385.

[185] A. L. Swindlehurst, A. K. Saxena, A. Mezghani, and I. Fijalkow, "Minimum Probability-of-Error Perturbation Precoding for the One-Bit Massive MIMO Downlink," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, 2017, pp. 6483-6487.

[186] C. Stockle, J. Munir, A. Mezghani, and J. A. Nossek, "Channel Estimation in Massive MIMO Systems Using 1-Bit Quantization," in *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Edinburgh, 2016, pp. 1-6.

[187] A. J. Roscoe, and R. A. Perrott, "Large Finite Array Analysis Using Infinite Array Data," *IEEE Trans. Ant. Propag.*, vol. 42, no. 7, pp. 983-992, July 1994.

[188] A. Ben-Tal, and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2001.